Moral Inferencing Patterns within Complex Post-traumatic Stress Disorder Populations

Helen Wang

Department of Psychology, Princeton University

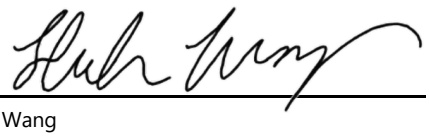
Dr. Yael Niv and Dr. Isabel Berwian

A Senior Thesis Submitted to Princeton University's Department of Psychology in Partial Fulfillment of the Requirements for the Degree of Bachelor of Arts

April 18th, 2022

Honor Pledge

I pledge my honor that this thesis represents my work in accordance to University regulations.

X 

Helen Wang

Acknowledgments

First and foremost, I want to thank my mother, who is, by far, the strongest woman I know. Thank you, Mom, for raising me even through all the hard times, for holding steadfast even when things were at their worst. Thank you for the sacrifices you've made to ensure that I was able to have a future. Thank you for the late hours you worked to make sure I had food in my belly and a roof over my head. Thank you for giving me the strength to make it as far as I have. I hope one day I'll make all the blood and tears you've shed worth it.

I want to thank Dr. Yael Niv, for opening this door in the first place. Thank you for your support and for your unrelenting kindness throughout this entire process. You made the lab feel like a home and your career (along with everything you fight for) will never not be an inspiration for all the women scientists following behind you. Thank you for everything!

As for Dr. Isabel Berwian, a simple 'thank you' seems insufficient. It's been a long two years but you've all but held my hand the entire way. Thank you for all the grace you've given me and for your never-ending patience. More than anything, thank you for always being willing to lend a listening ear and for guiding me throughout all of this. This project would not have been possible without you.

Thank you to the Niv Lab as a whole for always being willing to lend a hand. Special mentions to Branson Byers (who I bothered with many an IRB protocol issue) and Jamie Chin Chiu (for her assistance with Prolific).

I would also like to thank the people and institutions that made this thesis possible. Thank you to the Department of Psychology, the Program in Cognitive Science, and the Office of Undergraduate Research (as well as the Class of '55 Fund!) for funding this project in the first place. Thank you to Naila Rahman for her help with Qualtrics and thank you to Data & Statistical Services for always being willing to lend a hand. An especially big thank you to Dr. Jennifer Siegel, whose research served as the foundation for this thesis, for humoring all of my questions, as well as for her wonderful advice!

I absolutely cannot turn in my thesis without thanking all the friends and family that have accompanied me along the way. In no particular order: Thank you, Esme, for your never-ending passion and for sponsoring my late nights with your Keurig. Thank you, Iza, for your steadfastness and calm. Thank you, Emily, for your infectious cheer and consideration. Thank you, Leia, for being the girl-boss I inspire to be. Thank you, Breyonnah, for your companionship in all things, but especially for the solid two weeks where we camped at the library. Thank you, Olivia, for your grace in the face of adversity. Thank you, Jazz, for your persevering bravery. Thank you, Calvin, for being the only little brother I would ever want.

This thesis would not have been possible if not for the people around me. To that end, I am forever grateful.

Dedication

This thesis is dedicated to the woman who started me on this path in the first place, Dr. Bonnie A.B. Blackwell, who passed at the beginning of my senior year on September 8th, 2021. Dr. Blackwell, known affectionally as Dr. B or B^2 by her students, was a prolific scientist, known for her lifetime work in electron spin resonance dating. I knew her best as a high school teacher, whose no-nonsense demeanor pushed me farther than any other teacher before her. Without Dr. B, I would have never stepped foot into scientific research. Without Dr. B, I would never be where I am now. She opened many doors for many, many students and I was merely just one. For everything she's done, I'm forever grateful. Dr. B, your legacy is well felt, and I hope that, wherever you are, you are happy. We miss you. May you rest in peace.

Abstract

The validity of Complex Post-traumatic Stress Disorder (CPTSD) as an independent diagnostic category has long since been under question, due to its similarities to both Post-traumatic Stress Disorder (PTSD) and Borderline Personality Disorder (BPD). While research has been able to distinguish CPTSD from PTSD, the line between it and BPD is much less clear. This study focuses on a shared symptom cluster between the two disorders: deficits in interpersonal relations. Specifically, it attempts to examine moral inferencing patterns within CPTSD populations as past research has already been conducted assessing the mechanism in BPD populations. Using an adapted version of Siegel et al. (2018)'s moral inferencing task, where participants are asked to observe, predict, and rate the actions of two agents, one bad and one good, the study attempts to gauge for abnormal patterns. It further adds an additional element to the task by alternating the moral preferences of the two agents, to assess participants' abilities to respond to changing behaviors. Results found differences between the groups in their certainty ratings for the agents. In general, CPTSD group participants were more certain about their impression of the bad agent and less certain about the good agent. In contrast, the Control group was more certain about the good agent and less so about the bad agent. Interestingly, the trends observed in the moral inferencing patterns of the CPTSD group were more in line with that of individuals exposed to violence rather than that of BPD populations.

Keywords: Complex Post-traumatic Stress Disorder, Post-traumatic Stress Disorder, Borderline Personality Disorder, Disturbances of Self-Organization, Interpersonal deficits, Moral Inferencing.

Table of Contents

Introduction

The diagnostic validity of Complex Post-traumatic Stress Disorder (otherwise known as CPTSD or Complex PTSD) has been under contention since its development. While it was dismissed for inclusion within the Diagnostic and Statistical Manual for Mental Disorders (DSM) 5th Edition as an independent diagnostic category, it was incorporated as such within the 11th Edition of the International Classification of Diseases and Related Health Problems (ICD) (Cloitre et al., 2013). Critics of CPTSD have often questioned its validity, viewing it as redundant due to its similarities to both Post-traumatic Stress Disorder (PTSD) and Borderline Personality Disorder (BPD). Research has been able to establish clear differences distinguishing CPTSD from PTSD but has been far less successful in doing the same with CPTSD and BPD. The inability to clearly discriminate between the two has led many to speculate that CPTSD may actually be a byproduct of comorbid PTSD and BPD (Cloitre et al., 2013).

This study attempts to address the controversy by looking at interpersonal relationship deficits in CPTSD populations, as disturbances in interpersonal relations composes one of the main shared symptom clusters between CPTSD and BPD. As previous research has utilized moral inferencing paradigms when studying interpersonal relations within BPD, this study attempts to replicate the same paradigm within populations that reported CPTSD symptomology. It attempts to detangle interpersonal deficits within CPTSD populations by analyzing moral inferencing patterns with such populations when they are confronted with inconsistent moral preferences in others. In doing so, it also attempts to broadly assess a potential mechanism that may explain the patterns of interpersonal behavior exhibited while also situating the results reported in the current study within past moral inferencing literature with BPD.

Defining Complex Post-traumatic Stress Disorder

The concept of complex post-traumatic stress disorder, or CPTSD, was first proposed by Judith Herman in 1992, who noted a pattern of impaired self-organization amongst survivors of prolonged trauma that differed from classic post-traumatic stress disorder (PTSD) cases (Cloitre et al., 2013). She campaigned for the inclusion of an additional trauma-related disorder, one that she called Complex Post-traumatic Stress Disorder (Herman, 2015). Herman's proposed CPTSD was characterized by six distinctive symptom clusters: disturbances in affect regulation, alterations of consciousness, disturbed self-perception, disturbed perception of the perpetrator(s), interpersonal relationship difficulties, and changes in value systems (Maercker, 2021). Simultaneously, developmental psychologists were also arriving at a similar conclusion, with Bessel van der Kolk proposing what he coined as Developmental Trauma Disorder (DTD). DTD, unlike Herman's CPTSD, was specific to children and adolescents but addressed the same concerns as CPTSD, with it primarily focusing on the potential impact that multiple or chronic interpersonal trauma(s) could have on developing minds (Maercker, 2021). van der Kolk argued that such extensive trauma often resulted in dysregulation, altered attributions and expectations, and functional impairment.

Herman's CPTSD and van der Kolk's DTD were ultimately consolidated under the epithet "Disorders of Extreme Stress Not Otherwise Specified (DESNOS)" for the Diagnostic and Statistical Manual for Mental Disorders (DSM) IV field trials (Maercker, 2021). DESNOS largely utilized Herman's model, with its defining symptoms being affect dysregulation, dissociation, and somatization along with altered self-perception, relationships, and sustainable beliefs. Results from the field trials were promising, with the results indicating that participants with chronic trauma were more likely to report symptoms signifying disturbances in the

affective, self, and relational domains relative to participants with other traumas (Cloitre et al., 2013). Ultimately, however, it was rejected for inclusion within the DSM as an independent diagnosis. Instead, the DSM incorporated the additional symptoms found within the field trials under the broad umbrella term of PTSD.

The International Classification of Diseases and Related Health Problems (ICD), however, reached an alternative conclusion, electing to instead establish two sibling trauma disorders. They narrowed the definition of PTSD and added CPTSD as an independent diagnostic category (Maercker, 2021). As such, it should be noted, that within this study, the terms CPTSD and PTSD refer to ICD-11 definitions alone, not DSM definitions. As ICD-11 generates diagnoses exclusively based on symptom profile, trauma history is largely irrelevant in its clinical utility (though research has found clear differences) (Cloitre et al., 2013). As such, trauma history was not used as eligibility criteria for the study, though such information was collected for analysis.

As sibling disorders, PTSD and CPTSD share many similarities, including symptoms. Under the ICD-11, PTSD contains a set of six symptoms separated into three distinctive symptom clusters: (1) re-experiencing of traumatic events, (2) avoidance of trauma, and (3) sense of threat (Cloitre et al., 2013). To meet diagnostic criteria, individuals must experience at least one symptom within each cluster. CPTSD includes the aforementioned symptom clusters of PTSD as requirements for a formal diagnosis but is distinguishable from PTSD primarily through its additional three symptom clusters, which capture Disturbances in Self-Organization (DSO). (1) Affect dysregulation, (2) negative self-concept, and (3) interpersonal relational disturbance make up the DSO symptom clusters (Maercker, 2021).

Affect dysregulation refers to the tendency of CPTSD populations to swing between episodes of heightened emotional reactivity, violent outbursts, and self-destructive behavioral patterns to periods of prolonged dissociation (Cloitre et al., 2013). Emotional numbing and the inability to feel pleasure or positive emotions have also been reported. Negative self-concept, on the other hand, accounts for the persistent negative beliefs that plague individuals with CPTSD. They include beliefs about oneself being diminished, defeated, or worthless, as well as overwhelming feelings of shame and guilt. In the case of CPTSD, interpersonal relation difficulties present typically in the form of an inability to maintain relationships. CPTSD is largely marked by avoidance or lack of interest in social relations, often accompanied by the reported inability to feel close to others. In cases where relationships do develop, they are typically intense and individuals with CPTSD often report difficulties in maintaining emotional engagement.

Diagnostic Controversy and its Comparative Disorders

CPTSD has often been criticized for its similarities to both BPD and PTSD. Though not utilized as a criterion for clinical diagnosis, trauma type has repetitively shown up in research as a clear distinguishing factor between PTSD and CPTSD (Vang et al., 2019). Psychology typically acknowledges the existence of three types of traumas: Type I, Type II, and Type III (Luyten et al., 2019). Type I trauma refers to single-incident traumas, that are typically extreme and life-threatening but largely impersonal. Type II trauma addresses prolonged, chronic interpersonal trauma and Type III captures early life trauma that occurred within caregiving contexts. Type II and Type III traumas are often conflated into the broader category of Complex Trauma. The primary distinction is that unlike PTSD, which is more likely to result from Type I traumas, individuals with CPTSD are generally more likely to report experiencing traumas that

are (1) interpersonal, (2) of repeated and prolonged nature, (3) from which escape appeared difficult/impossible, (4) and typically (though not always) having occurred during childhood (Vang et al., 2019). In other words, CPTSD populations typically report higher rates of Type II and Type III traumas. Research has shown that each disorder has been associated with specific traumatic events that match such expectations. For instance, robbery is better correlated with PTSD while childhood sexual abuse is more strongly associated with CPTSD (Vang et al., 2019). Sexual abuse, domestic violence, torture, genocide, and childhood neglect/abuse, amongst others, are all forms of trauma that have been shown to lead to CPTSD (Luyten et al., 2019).

These findings were supported in Clotire et al. (2013)'s study, which evaluated whether CPTSD could be identified based on its clinical characteristics only. Using archival patient data from the New York City Trauma Clinic from 2002 to 2007, the researchers employed latent class analysis to identify groups. They found that a three-class model was significant and identified three classes that fit the labels of CPTSD, PTSD, and a low symptom group. 39.1% fit within the CPTSD group, 31.8% in the PTSD group, and 32.1% in the low symptom group. The three classes differed little demographically but did differ by trauma type. Logistic regression analyses found that reporting childhood abuse as the worst trauma was a significant predictor of CPTSD compared to PTSD, with those reporting childhood abuse as their worst trauma being nearly twice as likely to have CPTSD than PTSD. Conversely, those same analyses found that indicating single-event traumas like 9/11 as the worst trauma was significantly more predictive of PTSD than CPTSD. Odds ratio analysis signaled that those who reported 9/11 being their worst trauma were four times more likely to have PTSD than CPTSD.

Cloitre et al. (2019)'s population study echoed those findings, with their study noting that some traumas were associated with either CPTSD or PTSD specifically. The study looked at

non-institutionalized adults residing in the US that reported experiencing at least one traumatic event within their lifetime. A total of 1939 participants were examined and of those participants, 3.4% met criteria for PTSD while 3.8% met criteria for CPTSD. They found that cumulative childhood trauma was more strongly associated with CPTSD, though there were cases where adult trauma led to the disorder as well. Cloitre et al. (2019) also found that there were differences in disorder development depending on the trauma inflictor. Participants who experienced childhood sexual and/or physical abuse at the hands of a caregiver or guardian were more likely to develop CPTSD, for instance, while PTSD was more likely to develop if the assault was perpetrated by a non-parent/guardian.

Other distinctions between the two disorders have emerged within research. Individuals with CPTSD, for instance, generally reported lower rates of psychological well-being and higher scores for both Major Depressive Disorder (MDD) and Generalized Anxiety Disorder (GAD) than PTSD-afflicted individuals (Cloitre et al., 2019). Research generally suggests that CPTSD is the more debilitating of the two disorders, making up a higher portion of the clinical population and having greater reported functional impairment (Vasilopoulou et al., 2020). Furthermore, while PTSD has typically been construed as a fear condition, with symptoms often emerging as a result of contact with trauma-related stimuli, CPTSD's DSO symptoms, on the other hand, occur across a wide range of contexts and are not tied exclusively to trauma-related contexts (Cloitre et al., 2013). Vang et al. (2019) has also reported slight differences in memory recall between the two, having found that having a clear memory of the traumatic event(s) in question was positively correlated to PTSD symptomology but negatively correlated with DSO symptomology. Clearer recall was predictive of PTSD but not for CPTSD, suggesting deficits in

declarative memory for CPTSD populations, which conforms with reports of higher rates of dissociation with CPTSD versus PTSD (Vang et al., 2019).

Comparative Symptomology and Presentation

To meet diagnostic criteria for CPTSD, individuals are required to meet symptom criteria for PTSD while also endorsing symptoms from each of the DSO symptom clusters. Notably, CPTSD's DSO symptom clusters heavily overlap with the symptomology of BPD (Cloitre et al., 2013). BPD is a psychiatric disorder characterized by impairments in one's sense of self and identity, with symptoms including dissociation, impulsivity, problematic interpersonal relationships, emotional dysregulation, and hypersensitivity to social exclusion (Luyten et al., 2019). The boundary between BPD and CPTSD remains murky. Much of the argument that BPD and CPTSD are related is based on the mechanisms behind the disorder and their similar symptomology. Both disorders are hypothesized to be a result of maladaptive relationships with caretakers, and both are characterized by their tendency to replicate those relationships with others in adulthood (Laddis, 2010). They are also equally characterized by guilt, shame, loss of faith in others, hopelessness, and mistrust (Laddis, 2010). High comorbidity rates between PTSD and BPD also have many suggesting that CPTSD is merely a product of this comorbidity. Within the US, researchers have found that 24% of individuals with PTSD have comorbid BPD while 30% of individuals with BPD meet the criteria for PTSD (Cloitre et al., 2014). The rates are even greater in clinical samples, with BPD comorbidity for PTSD patients ranging from 37 to 68% and PTSD comorbidity for BPD patients running as high as 58% (Cloitre et al., 2014).

However, some researchers have pointed out fundamental distinctions in how symptoms in the two disorders are presented. Both disorders have issues with self-concept but whereas CPTSD is characterized by persistent negative self-concept, BPD's primary issue revolves

around an unstable sense of self (Hyland et al., 2019). Relational difficulties plague the two equally but patients with BPD tend to experience volatile relationships characterized by constantly switches between idealization and devaluation, coupled by a persistent fear of abandonment (Cloitre et al., 2014; Hyland et al., 2019). The interpersonal deficits of CPTSD populations, however, are better characterized as the chronic avoidance of relationships (Cloitre et al., 2014; Hyland et al., 2019). Furthermore, emotional dysregulation is fueled by a fear of abandonment for BPD rather than a struggle with modulating it, as it is for CPTSD (Hyland et al., 2019). Emotional sensitivity, reactive anger, and poor coping responses are typically more characteristic of CPTSD. While the aforementioned also occur within BPD populations, affect dysregulation typically manifests itself in BPD in the form of high suicidality and self-injurious behavior, which is noticeably less prominent in CPTSD (Cloitre et al., 2014). Roughly 50% of individuals with BPD report such behaviors whilst only 14.3% and 16.7% of CPTSD and PTSD individuals do so (Cloitre et al., 2014). More importantly, however, a key point of differentiation is that the presence of a traumatic stressor is not required for BPD like it is for CPTSD (Hyland et al., 2019). While indeed, many BPD patients do report high rates of trauma, the literature has also indicated that anywhere between 8% to 70% of BPD patients do not report any history of early abuse or neglect at all (Luyten et al., 2019).

Cloitre et al. (2014) conducted a latent class analysis with an archival dataset of 280 female patients with a history of child abuse to determine if separate classes for PTSD, CPTSD, and BPD could be detected. The latent class analysis produced a four-class model that covered PTSD, CPTSD, BPD, and a low symptom group that had acceptable levels of discrimination. Based on symptomology, the study found that only 7.8% of the CPTSD class met the criteria for BPD compared to the 91.9% of the BPD class. Similarly, only about 44.6% of the BPD class met

criteria for CPTSD. Furthermore, they found that BPD class individuals were actually more likely to meet the criteria for PTSD (at 54.9%) rather than CPTSD (at 45.1%). The study identified four symptoms that predicted classification into BPD rather than CPTSD: (1) efforts to avoid abandonment, (2) unstable relationships characterized by oscillation between idealization and devaluation, (3) unstable sense of self, and (4) impulsiveness. They also found that BPD class individuals were more likely to endorse suicidal or self-injurious behaviors at 50%, versus the 14.3% and 16.7% of the CPTSD and PTSD classes. BPD and CPTSD class individuals both were found, however, to have similarly high levels of functional impairment when compared to PTSD and low symptom classes.

Theory of Development

Research focused on CPTSD has largely examined it from the perspective of childhood trauma, particularly through the lens of attachment style in abused children. Bowlby's attachment theory, for instance, theorizes that abused children are able to preserve their attachment to their caregivers via dissociation. The theory proposes that such children form idealized attachments to their abusive and/or neglectful caregivers while simultaneously dissociating or denying negative interactions (Zamir & Lavee, 2014; Zulueta et al., 2019). These children have insecure or disorganized attachment, with the children often alternating between clinging to and avoiding caretakers (Nieuwenhove & Meganck, 2019). This maladaptive relationship between child and abusive caretaker serves as the child's fundamental working model for relationships, which they utilize in their later life to disastrous effects within their social circles. While such models may have been advantageous within the specific abusive environments, they often function poorly outside of it.

In healthy caregiver-child relationships, children learn adaptive emotion regulation and coping skills by using caregivers as support and as a model (Nieuwenhove & Meganck, 2019). Without a secure attachment to caregivers, however, abused children lack templates for proper interpersonal and individual development (Nieuwenhove & Meganck, 2019). For instance, because they employ avoidance tactics such as dissociation or denial to cope with their environment, individuals with complex trauma typically lack emotional awareness and struggle with identifying dangerous interpersonal contexts (Zamir & Lavee, 2014). There are reports of alexithymia, or inability to voice their own emotions, within CPTSD populations that may hinder their ability to share and cope (Zulueta et al., 2019). Their difficulties with coping with intense emotions such as rage or terror, often lead them to resort to maladaptive, self-soothing behaviors such as substance abuse, self-injury, and/or risky sexual behavior (Pearlman & Courtois, 2005). As a result of their impaired emotional regulation, CPTSD populations have repeatedly been linked with high comorbidity to alcohol and substance abuse disorders (Hyland et al., 2018).

Perhaps even more worrisome is the impact that complex trauma has on an individual's self-concept. Individuals with CPTSD are typically marked by persistent negative self-concept, conceptualizing themselves as weak, worthless, and defeated, beliefs that are often accompanied by shame or guilt (Zulueta et al., 2019). There is also a trend of self-blame, followed by the persistent idea that they are perhaps more deserving of abuse or pain than others (Nieuwenhove & Meganck, 2019). Such self-conceptualizations, unfortunately, come hand-in-hand with poor interpersonal relations, another defining symptom cluster for CPTSD.

Patients with CPTSD often report difficulties in interpersonal relationships. They typically view others with mistrust and suspicion, viewing them as dangerous and unpredictable (Nieuwenhove & Meganck, 2019). This often leads them to persistently avoid relationships when

possible (Hyland et al., 2019). For instance, people with complex trauma often struggle with the concept of relying on others (Luyten et al., 2019). However, they also tend to swing between avoiding relationships and desiring them, much like the insecure/disorganized attachment that characterized their maladaptive relationships with previous abusers (Nieuwenhove & Meganck, 2019). Oftentimes when individuals with CPTSD do form relationships, they are intense and close ones that they struggle to maintain (Cloitre et al., 2013). An alarming trend amongst those with CPTSD is their tendency to unconsciously replicate maladaptive relationships. They often repeat the same power dynamics of dependency while simultaneously also struggling with communicating and negotiating relationship boundaries (Nieuwenhove & Meganck, 2019). Simultaneously, their disorganized attachment often manifests in forms such as traumatic bonding with the malignant, parentification/caretaking of others, and extreme dependency that leaves them vulnerable to exploitation (Pearlman & Courtois, 2005). Consequently, there are often high rates of revictimization amongst victims of complex trauma, with early childhood trauma being continuously linked with later revictimization in adulthood.

Moral Inferencing Literature

Interpersonal difficulties characterize both BPD and CPTSD populations. One important component of interpersonal interaction is the ability to form and maintain accurate representations of others, specifically their moral character (Siegel et al., 2020). As a main component of moral character, the ability to assess the harmfulness of others is critical for viable interpersonal interactions. The ability to do so is typically contingent on two components: (1) an individual's capability of updating beliefs objectively so as to predict future outcomes and (2) an individual's ability to evaluate and form impressions about another's moral character (Siegel et al., 2019). The latter is important in terms of determining whether forming a relationship with an

individual is in one's interest while the former is important with regard to deciding whether maintaining such a relationship is beneficial. This ability is broadly referred to within this paper as one's capacity for moral inferencing.

In general, studies have shown that environmental experiences such as exposure to violence can impact this learning. Siegel et al. (2019) conducted a study evaluating how violence exposure can impact harmful/harmless moral character learning in 119 incarcerated males. Participants were asked to predict and observe the choices of two agents, one that was programmed to behave like a good agent and another programmed to exhibit the behavior of a bad agent, both of whom were asked to inflict shocks on a third party individual for varying sums of money. After predicting the choices made by the agents, participants were given feedback on the accuracy of their assessment. Based on those observations and the feedback they received, participants were then asked to rate their impression of the agent's moral character as well as the certainty of their impression. The study was then concluded with a one-shot trust game.

Siegel et al. (2019) found that the participants were equally accurate in their predictions for both agents at 72% for the good agent and 77% for the bad agent, suggesting that they were equally motivated to learn about the harm preferences for both. The study found that exposure to violence did not necessarily impact harm learning in the sense that it didn't affect belief updating rates, but it did impact impression formation. Participants that had high exposure to violence were more likely to rate both agents more negatively when compared to control groups. In terms of overall impressions, however, exposure to violence generally led to more favorable impressions of the bad agent and less favorable impressions of the good agent, which indicated that there was little differentiation between the two agent's moral characters. Exposure to

violence did not make participants more likely to predict worse intentions prior to the onset of the task and they were not less trusting of others. The authors did note, however, that there were differences in uncertainty ratings, observing that participants that were exposed to violence were more uncertain that the good agent was good but less uncertain about whether the bad agent was bad. Such uncertainty was replicated in the concluding trust game, with participants displaying maladaptive trusting behavior when interacting with good agents, in that they entrusted them with less money than healthy controls. Such results are interesting in the context of CPTSD, as individuals diagnosed with such a disorder are far more likely to be exposed to such violent experiences.

Another study conducted by Abramov et al. (2020) examined trust behavior in BPD in closer detail using a multi-round trust game. They attempted to study how BPD traits could influence responses to trust violations and efforts for trust reparation, more specifically the direction and rates of change of trust. Abramov et al. (2020) evaluated 234 undergraduate students for BPD traits using the McLean Screening Instrument for Borderline Personality Disorder. Based on the screening, they formed a low-BPD group, a Moderate-BPD group, as well as a High-BPD group that met diagnostic criteria for BPD. Participants engaged in a 15-round trust game where they were experimentally exposed to a trust violation via negative changes in the rate of reciprocity after the fifth round for three rounds. The game was then rigged to simulated behavior that would induce trust repair by changing the rate of reciprocity to a more favorable one after the ninth round. This was done to stimulate trust formation, dissolution, and restoration. For the study, they used the amount of money invested as a measure of trust.

In general, Abramov et al. (2020) found that the number of BPD traits affected the rate of initial trust growth, with those with higher BPD traits generally being slower to form trust, as

indicated by less money invested. What was unexpected, however, was that BPD traits predicted increased trust growth following a trust violation. High BPD groups invested more money following such an event. After trust reparation began, however, high BPD participants notably decreased the money invested. The authors hypothesized that the reparative behavior on the part of the trustee may have triggered caution and suspicion in high BPD trait individuals. They noted that high BPD trait individuals appeared wary when confronted with cooperative behavior and become contradictorily more trusting when treated inequitably. The implications of the study are alarming when placed in the context of interpersonal relationships, where such tendencies may quite well lead to maladaptive relationships.

Some have speculated that deficits in such areas contribute to the tendency of BPD patients to struggle with maintaining interpersonal relationships. The ability to form accurate representations of others' characters and intentions is crucial to social learning; deficits in such areas can have long-term impacts on their quality of life (Siegel et al., 2020). Similar perception difficulties have been reported before amongst the BPD population. Notably, BPD individuals have exhibited a bias toward perceiving neutral faces as negative (Luyten et al., 2019). Other studies have also noted that BPD, as well as CPTSD, has exhibited biases in encoding negative memories over positive memories (Ford & Courtois, 2014).

Within CPTSD, some research has supported the notion that interpersonal disturbances are partially caused by emotional processing biases that result from affect dysregulation. Bertó et al. (2017) looked at emotional processing within maltreated children diagnosed with CPTSD. Using a visual dot-probe task, a behavior task designed to examine relationships between attention and emotional stimuli, they assessed 47 children, with 21 of the children having experienced maltreatment and matching symptom criteria for CPTSD. They found clear

differences between their experimental and control groups, with the CPTSD group exhibiting attentional biases away from angry faces and attentional biases towards sad faces. The avoidance of angry faces suggested that CPTSD individuals may have a higher threshold for detecting threats, which may leave them vulnerable to dangerous interpersonal interactions.

Siegel et al. (2020) recently investigated moral inferencing in patients with BPD as well as with patients with BPD treated with democratic therapeutic community (DTC) treatment. They replicated the same methodology as Siegel et al. (2019) with a population of 43 participants with diagnosed BPD (both treated and untreated) and 106 controls. They found that there were no differences in initial character impression but that there were differences in uncertainty. Patients with BPD were more certain about bad agents and were more uncertain about good agents. They were generally slower to update beliefs about the bad agent but were faster to update beliefs about the good agent. DTC-treated BPD participants, when compared to untreated patients, were more likely to have more favorable impressions for the good agent. DTC-treated patients were also more uncertain about the bad agent and had faster learning rates but held similar levels of uncertainty and learning rates for the good agent when compared to untreated patients. The study pointed to the tendency of BPD patients to form more rigid impressions about bad agents than good agents. While there is an abundance of data assessing mechanistic causes behind interpersonal deficits and maladaptive relationship formation for BPD, there is little to none examining the same construct for CPTSD, despite the controversies surrounding the two disorders.

Objective and Hypotheses

Considering the similarities between CPTSD and BPD as well as controversy over the diagnostic validity of CPTSD, the current study aims to examine interpersonal deficits within the

CPTSD population. As both CPTSD and BPD are marked by difficulties in interpersonal relationships, which have often led to devastating rates of revictimization for both groups, examining the mechanisms behind interpersonal disturbances in CPTSD populations may provide some insight into potential similarities and differences between the two disorders. Researchers have traced the issue of moral inferencing as a potential explanation for disturbed interpersonal relationships in BPD. This study aims to examine how moral inferencing works in CPTSD and how it may or may not differ from how it presents within healthy controls. It also further attempts to evaluate moral inferencing in individuals who have experienced childhood trauma but do not exhibit CPTSD symptoms in an attempt to assess whether impaired moral inferencing is a result of CPTSD alone or merely a byproduct of an extensive trauma history.

To do so, the study utilizes an adapted variation of Siegel et al. (2020)'s moral inferencing task that has three phases: impression formation, impression violation, and impression reparation. In the first phase, participants form initial impressions of a good and bad agent that are then later violated in the second phase. The third phase involves realigning behavior to match that of the initial impression. It aims to closely examine participants' initial moral character impression formation, their level of uncertainty or certainty of those impressions, and their ability to accurately predict actions based on those impressions. In doing so, the study attempts to evaluate the mechanisms behind interpersonal deficits for CPTSD.

Hypothesis 1

Patients with CPTSD hold a general sense of mistrust for all individuals. Regardless of whether individuals exhibit "good" or "bad" behavior, individuals with CPTSD may view them equally. Therefore, it is predicted that differences in impression ratings between good versus bad agents would be minimal for individuals with CPTSD relative to the other groups.

Hypothesis 2

As one characteristic of interpersonal relationships in individuals with CPTSD is the tendency to mistrust, and therefore avoid, others, it follows that individuals with CPTSD may be equally suspicious of all people. Positive information about others may be met with suspicion while negative information may merely validate the mistrust. As such, it is predicted that individuals with CPTSD, upon forming an impression, will display similar amounts of certainty for both good and bad agents, such that differences in certainty ratings between the two agents would be minimal compared to individuals with past childhood trauma or healthy controls.

Hypothesis 3

Patients with CPTSD tend to have high victimization rates for maladaptive relationships. A potential explanation could be that people with CPTSD, after forming an impression, are slow or resistant to altering that impression. Such an explanation could explain why people with CPTSD continue to stay in maladaptive relationships even as abusive behaviors emerge. Therefore, it is predicted that after the impression formation phase, individuals with CPTSD will display smaller changes in mean impression rating when compared to other groups.

Hypothesis 4

If Hypothesis 3 is correct, then participants with CPTSD will be resistant to any contradicting information they may encounter after the impression formation phase. As such, it can be assumed that after impression formation, participants with CPTSD may have lower prediction accuracy than those without as they would not take the new information into account in their predictions.

Methodology

Recruitment

Participants were recruited through Prolific, an online recruitment platform, and then redirected to a Qualtrics survey for the study. Recruitment occurred in 2 phases, with participants being first invited to complete a prescreening survey, upon which they would receive an invitation to complete the main survey if they met criteria for one of the arms of the study (see below for criteria). The study was restricted to participants who were over 18 years of age and to individuals that indicated that they were located within the United States. Participants who did not complete the entire study were excluded from data analysis. Study completion, in this case, was defined in the study as completion of the prescreen measures, the adapted moral inferencing task, and demographic questions (see Figure 1). To ensure data quality, participants that were flagged as potential bots by Qualtrics were also not included. Similarly, participants who failed the infrequency attention checks were also excluded (see Appendix A).

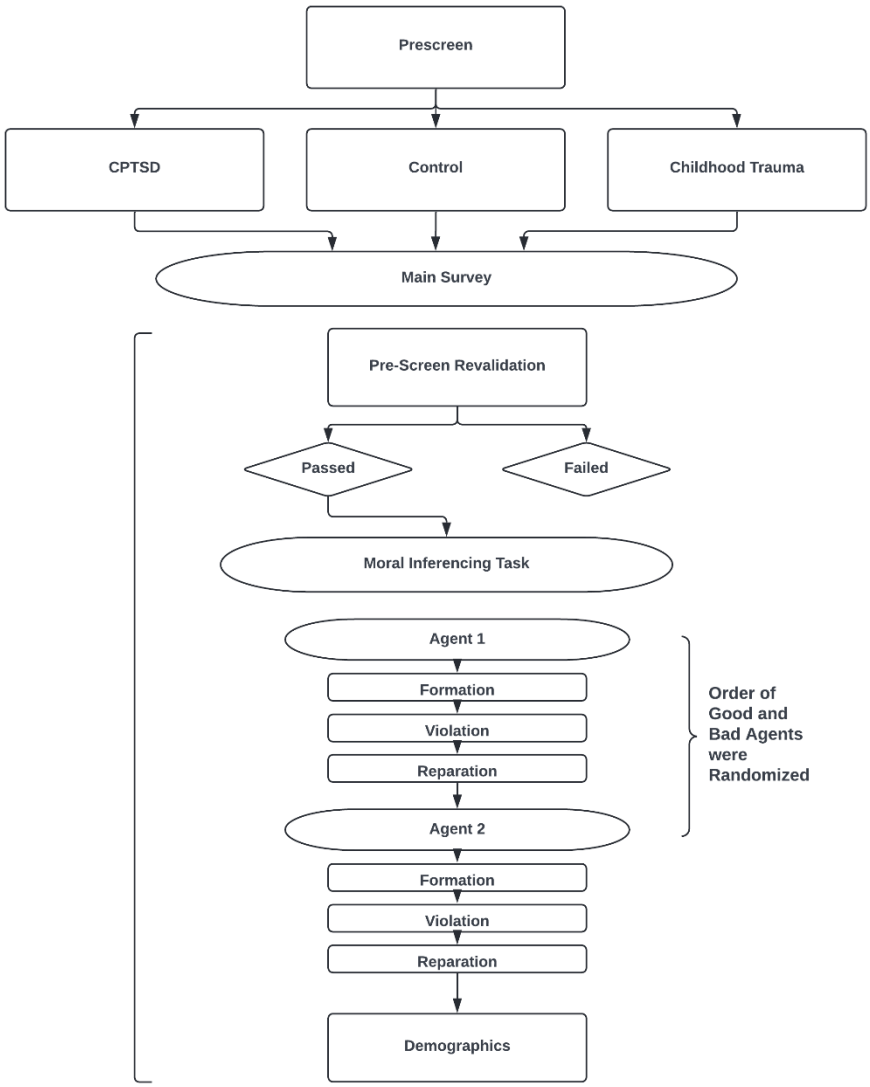
Participants were compensated at a rate of $12/hour, with participants being given $1 for completing the estimated 5 minute prescreen and $8 for completing the estimated 40 minute main survey. Bonus payment was given for the main survey, with a maximum additional allotment of $1.80, to incentivize participants to answer attentively. Bonus payment allotment was calculated based on task performance, at a rate of $0.01 per correct answer. Partial payment was given to participants who failed to complete the survey at a rate that matched the percentage of the survey they completed. Participants who were invited but failed to re-meet criteria eligibility in the main survey were eliminated from the survey but compensated at a flat rate of $0.41 (corresponding to the percentage of the survey they completed) for the main survey.

Demographics

Participants were further asked to report information about their racial-ethnic status, highest level of education, gender, financial status, marital status, and age. Frequency and percentages were reported for each category were reported for each group.

Figure 1

Experimental Design and Survey Flow



Prescreen

The study consisted of three groups: CPTSD, Childhood Trauma, and Control. The CPTSD group consisted of individuals who met the diagnostic criteria for CPTSD per ICD-11 criteria, as measured by the International Trauma Questionnaire (ITQ). The Childhood Trauma group consisted of people who reported no CPTSD symptomology but had a history of childhood trauma, as assessed through the Childhood Trauma Questionnaire – Short Form (CTQ). This group was created to help determine whether any patterns observed could be attributed solely to CPTSD or were merely a broader result of childhood trauma as a whole. Participants that met criteria for the Childhood Trauma group generally reported no clinical symptoms (scores < 2) on any of the diagnostic questions in the ITQ and reported at least one 'Severe' or two 'Moderates' for the CTQ subscales. The Control group, on the other hand, consisted of participants who reported no symptoms of CPTSD and reported none or low amounts of Childhood Trauma. Participants were eligible for the control group if they reported no clinical symptoms (scores < 2) on any of the diagnostic questions in the ITQ and did not receive 'Severes' or 'Moderates' on any of the CTQ-SF subscales. There is currently no scoring system available for the total CTQ score. As such, the study elected to use subscale scores rather than the total CTQ score for group eligibility criteria.

International Trauma Questionnaire (ITQ)

The International Trauma Questionnaire (ITQ) is a brief self-report measure created by Cloitre et al. (2018) to evaluate for PTSD and CPTSD symptomology using ICD-11 diagnostic criteria. It contains twelve symptom indicators, evaluated on a 5 point likert scale ranging from 'not at all' (0) to 'extremely' (4). The ITQ considers symptoms clinical present if a score of

‘moderately’ (2) or higher is indicated. The first six items measure PTSD, two each per each PTSD symptom cluster (reexperiencing threat, avoidance, and sense of threat). The latter six measure Disturbances in Self-Organization (DSO) symptoms, CPTSD’s additional cluster. Two items each are used to measure the main symptoms of the cluster (affective dysregulation, negative self-concept, and disturbances in relationship). For the two items evaluating affective dysregulation, one measures hyperactivation while the other measures hypoactivation. Another six items are also included that evaluate for functional impairment in relation to PTSD symptoms and DSO symptoms. To meet criteria for CPTSD, participants must endorse at least one symptom from each PTSD and DSO symptom cluster while also indicating a clinical level of functional impairment for both PTSD and DSO Sections. The ITQ exhibited excellent internal consistency within the study for both PTSD (α =.913, 95% CI [.900, .922]) and DSO items (α = .908, 95% CI [.896, .918]) as well as a whole (α = .942, 95% CI [.934, .948]).

Childhood Trauma Questionnaire – Short Form (CTQ)

The Childhood Trauma Questionnaire – Short Form (CTQ) is a self-report questionnaire containing 28 items used to assess interpersonal childhood trauma. It was created by Bernstein et al. (2003) and is a condensed variation of the original 70 item Childhood Trauma Questionnaire (CTQ). Items are scored on a 5 point Likert scale, ranging from ‘never true’ (1) to ‘very often true’ (5). Items can be scored to produce five subscales that measure emotional abuse, physical abuse, sexual abuse, emotional neglect, and neglect. An additional subscale measures minimization/denial of traumatic events and captures underreporting. Subscale scores were used to further grade level of trauma in each category on a scale ranging from ‘None’ to ‘Low’ to ‘Moderate’ to ‘Severe’ as based on Bernstein & Fink (1998). The CTQ exhibited excellent to good internal consistency within the study as a whole (α = .893, 95% CI [.878, .906]) as well as

for its emotional abuse (α = .899, 95% CI [.883, .913]), physical abuse (α = .849, 95% CI [.816, .875]), sexual abuse (α = .951, 95% CI [.939, .960]), emotional neglect (α = .931, 95% CI [.919, .940]), physical neglect (α = .809, 95% CI [.777, .836]), and minimization/denial subscales (α = .900, 95% CI [.883, .915]).

Adverse Childhood Experiences scale (ACEs)

The Adverse Childhood Experiences scale (ACEs) is a self-report measure created by Wade et al. (2017) that evaluates for negative experiences in childhood. It contains 10 items that evaluate for various forms of childhood abuse, including emotional abuse, physical abuse, sexual abuse, and physical neglect, as well as general household dysfunction. In the latter case, it evaluates for exposure to substance abuse, mental illness, violent treatment of maternal figures, and criminal behavior within the household. All questions are scored on a binary basis, ranging from 'yes' (1) to 'no' (0), with the questionnaire itself exhibiting good internal consistency (α = .805, 95% CI [.778, .827]).

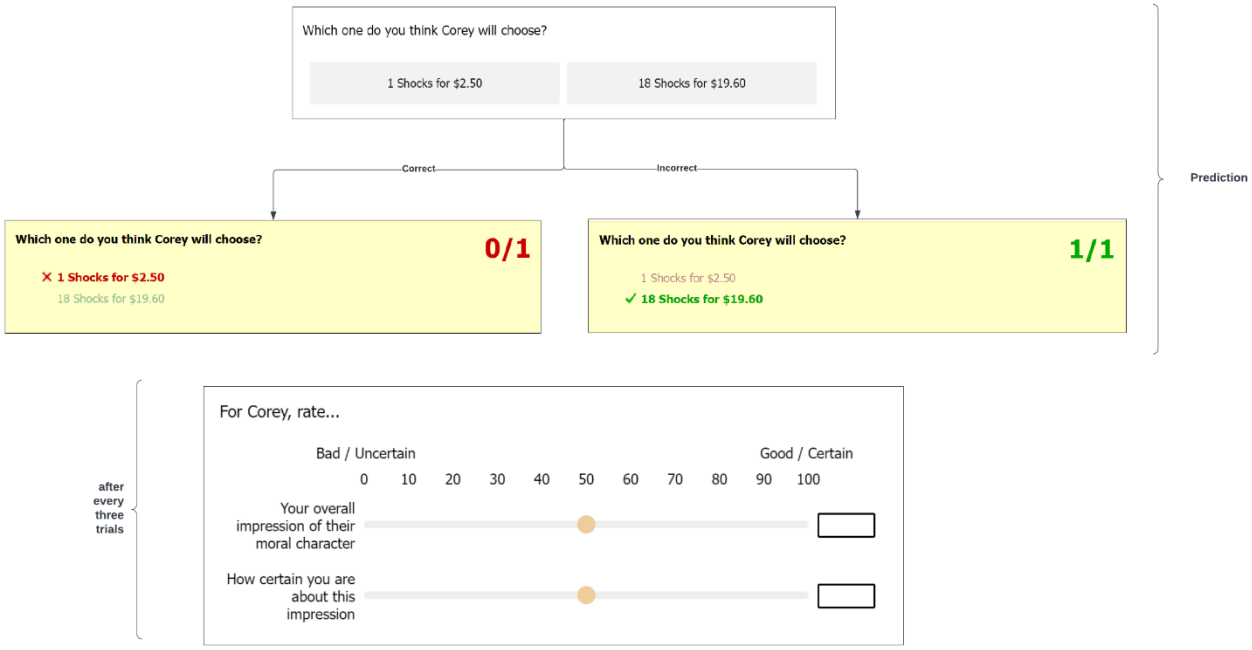
Adapted Moral Inferencing Task

Upon completing the prescreen, participants were introduced to an adapted form of Siegel et al. (2020)'s moral inferencing task. In the original task, participants observed and predicted a series of fifty choices made by two agents, receiving feedback on their predictions immediately after each trial. Participants observed the entire sequence for one agent before observing the next. The order of the agents was randomized. In these series of choices, agents were given two options, one where they received more money in exchange for more electric shocks to a third party victim (C_{harm}) and another where they received less money in return for fewer shocks to the victim (C_{help}) (see Figure 2). For the task, no prior information was provided

on the agent's moral character. Rather, the task was designed to allow the participant to learn about the agent's moral preferences through trial and error.

Figure 2

Adapted Moral Inferencing Task



Within Siegel et al. (2020)'s design, the agents had varying moral preferences, with one exhibiting the behavior of a 'bad' agent and the other exhibiting behavior of a 'good agent.' These moral preferences stayed consistent throughout the task. After every third trial, participants were also asked to report their impression of the agent's moral character (from 0 = 'bad' to 100 = 'good') as well as their certainty of the impression (from 0 = 'uncertain' to 100 = 'certain'). To minimize the influence of prior expectations about people's general moral character, Siegel et al. (2020) anchored prior expectations by telling participants explicitly that people typically required $1 per shock to the victim. Before the onset of the task, Siegel et al.

(2020) also measured participants' prior expectations by asking them to report their expectations about the overall moral character of the agents.

In the adapted form of Siegel et al. (2020)'s task, the general framework of the task was retained. However, the adapted form utilizes 90 trials per agent rather than 50 and is altered such that the agent's moral preferences change throughout the 3 stages of the trials (see Table 1). Each stage consisted of 30 trials. In the first stage, the Impression Formation stage, the agent 'Alex' exhibits choice behavior along the lines of a 'good' agent whilst the agent 'Corey' does the opposite, displaying choice behavior more characteristic of a 'bad' agent. In the second stage, the Impression Violation stage, the moral preferences of the agents switch, with 'Corey' behaving like a 'good' agent and 'Alex' behaving like a 'bad' agent. In the final stage, the agents return to their original moral preferences. The stages are designed to assess how moral inferencing accounts for changes in moral character, a characteristic of maladaptive relationships. For the duration of the study, the original and final moral preferences will be used to indicate the agents, i.e. Alex will be referred to as the good agent for the duration of this paper. Like Siegel et al. (2020), the study also anchored the prior expectations of participants in the same manner and also asked participants to report their expectations of the overall moral character of the agents.

Table 1

Moral Characters of the Agents in Each Task Phase

Agent	Task Phase		
	Impression Formation	Impression Violation	Impression Reparation
Alex ("Good")	Good	Bad	Good
Corey ("Bad")	Bad	Good	Bad

Generating Moral Preferences

Siegel et al. (2020)'s methodology was utilized to generate the choices and stimulate decisions made by the two agents, Alex and Corey, such that they matched their assigned moral character (see Appendix B). Supplementary Information from Siegel et al. (2018) and Siegel's MATLAB code was consulted to generate moral preferences (J. Siegel, personal communication, November 10, 2021). Moral preferences, or preferences towards inflicting harm on the third-party victim, are operationalized in the study as the exchange rate between money for the agents and shocks for the victim, otherwise known as the harm aversion parameter κ. When $\kappa = 0$, agents are willing to accept any number of shocks for the third-party victim in exchange for money. As κ approaches 1, agents are more harm averse, becoming more willing to reject large quantities of money in exchange for not shocking the victim. For the moral inferencing task, κ was used to distinguish between agents that were 'good' ($\kappa = 0.7$) versus 'bad' ($\kappa = 0.3$). As such, in terms of decisions, a 'bad' agent was more willing to choose options that would inflict harm on a third-party victim if given monetary incentives compared to a 'good' agent. Though both agents would be presented with identical sets of options, their differing moral preferences would lead them to choose differently in many cases.

Each trial consisted of two choices [s-, m-] and [s+, m+] that matched a specific κ value. [s-, m-] consist of the option that the agent would pick if they chose the choice with the least shocks to the victim (C_{help}) while [s+, m+] was chosen if the agent chose the choice with more shocks to the victim (C_{harm}). To generate the choices, for each phase, 15 values were randomly drawn from a normal distribution around the 'good' agent's κ value ($M = 0.7$, $SD = 0.15$), where $\kappa < 0.95$. From those values, a matched set of κ values were generated for the bad agent by subtracting each newly drawn κ value from 1. This was done to ensure that participants would

receive equivalent amounts of informational value for the good agent as well as the bad agent. As such, each phase contained 15 trials that were informational about the 'good' agent's κ value and 15 trials that were informational about the 'bad' agent's κ value. The order of these trials was then randomized within each phase. As the adapted task contained three phases, that left each agent with 90 trials, each containing a pair of choices. For the first and last trials in the task, the κ value was fixed such that $\kappa = 0.5$.

Shock and money options were then generated using the κ values. For each κ value, 10000 random positive shock differences (Δs) were generated from a uniform distribution such that $1 < \Delta s < 20$. Δm, or positive money differences, was then derived using the following equation such that $0.10 < \Delta m < 19.90$:

$$\Delta m = \frac{\kappa_t \Delta s}{1 - \kappa_t} \quad (1)$$

Where κ_t refers to the generated κ value used for trial t. The [Δs, Δm] pair was then used to generate the choice options [s-, m-] and [s+, m+] for each trial t. s- was drawn from a uniform distribution of positive integers ranging from 0 to 20 while m- was drawn from a uniform distribution of positive numbers running from 0 to 20, rounded to the nearest 10th. s- and m- were constrained such that $0 < s- + \Delta s < 20$ and $0 < m- + \Delta m < 20$ respectively. s+ and m+ were then calculated by adding Δs and Δm to s- and m-.

To stimulate the agents' decision, V_{harm}, or the utility of choosing the more harmful option, was then calculated using the following equation:

$$V_{harm} = (1 - \kappa_n)\Delta m - \kappa_n \Delta s \quad (2)$$

Where κ_n refers to the κ for agent n in that phase ($\kappa_{bad} = 0.3$, $\kappa_{good} = 0.7$).

To then calculate P_{harm}, or the probability of an agent choosing the more harmful option (C_{harm}), the following equation was utilized:

$$P_{harm} = \frac{1}{1 + e^{-\beta \times V_{harm}}} \quad (3)$$

Where β defines how deterministic an agent's choices are. When β approaches 0, there is a significant amount of noise in the agent's choice whereas when β approaches 100, agents were more deterministic in their choice preferences. For the study, β was fixed to 100 to stimulate agents that made choices in a deterministic fashion.

P_{harm} was transformed into a binary choice, u, using the following Boolean equation:

$$u = \begin{cases} 1, & x_{rand} < P_{harm} \\ 0, & x_{rand} \geq P_{harm} \end{cases} \quad (4)$$

where x_{rand} is a random number between 0 and 1. If u = 1, the agent chose C_{harm}, otherwise known as the choice with the most shocks. If u = 0, or false, the agent would choose C_{help}, the option with the least shocks.

Data Analysis

Analyses were conducted using RStudio (R version 3.6.1) (R Core Team, 2019). To assess demographical differences between groups, Fisher's Exact Test was conducted for race, education, gender, financial status, age, and relationship status variables. Pearson's Chi-square Test was not used as expected values were less than 5 and was therefore inappropriate for the current analysis. Non-parametric Mann-Whitney U Tests were conducted to assess for differences in total ACES score, total CTQ score, and CTQ subscale scores due to failure to meet assumptions of normality. As there were tied values within the data, the wilcox_test() from the 'coin' package was utilized to conduct the Mann Whitney U test rather than the standard

wilcox.test() R function (Hothorn et al., 2016). The wilcox_test() within ‘coin’ is built to be able to calculate exact p-values even in the presence of ties, which the wilcox.test() in R is unable to do.

Data for impression ratings, certainty ratings, and prediction accuracy were aggregated for ANOVA analysis, such that the mean score for each Group x Agent x Phase interaction for each subject was calculated. For certainty and impression, rating scores were utilized. For prediction accuracy, as scores were graded 1 (Correct) and 0 (Incorrect), scores were averaged to get composite scores. The data was analyzed to assess if it met assumptions for ANOVA using the Shapiro-Wilk Test for Normality and Levene Test for Homogeneity of variance via the ‘rstatix’ package (Alboudkadel, 2021).

If assumptions were met, a three-way Mixed ANOVA was run (Singmann et al., 2021). Mauchly’s Test of Sphericity was conducted, and Greenhouse Geisser Sphericity corrections were applied as needed. If assumptions were not met, the non-parametric alternative, the Aligned Rank Transformation ANOVA (ART ANOVA), was conducted using the ‘ARTool’ package instead (Kay et al., 2021).

For the Mixed ANOVA, post hoc analysis was conducted by applying contrast tests via pairwise comparisons, with Bonferroni adjustment, through the ‘emmeans’ package (Russell, 2022). For the ART ANOVA, contrast tests using pairwise comparisons were performed using the ART procedure via the art.con() function, with p-values adjusted using the Bonferroni method. Line and bar plots generated to showcase the analyses were created using the ‘ggplot2’ package (Wickham, 2016). Power analysis was conducted using G*Power software (Faul et al., 2007).

Results

Participants

A total of 650 participants completed the initial pre-screening. Of those 650, 9 were excluded for being flagged as bots by Qualtrics's Bot Detection service. Of the remaining 641 participants, 89 of the participants met criteria for CPTSD in the prescreen and were re-invited back to complete the main survey. Of the 89, 70 participants attempted to complete the main survey. Of the 70, 28 failed to re-qualify in the prescreen revalidation, 4 failed to complete the entire survey, and 2 were flagged as bots. Only 37 participants in the CPTSD group ultimately re-qualified in the prescreen re-validation, completed the entire survey, and were not flagged as bots. A further 2 participants of the CPTSD group were excluded for failing the attention check, leaving a total of 35 participants for that group.

A total of 142 and 37 participants met criteria for the Control and Childhood Trauma respectively from the pre-screen. 59 participants responded to the invite for the Control group. Of the 59, 7 failed to complete the entire survey, 2 were flagged as bots, and 12 failed to requalify in the prescreen validation. For the Childhood Trauma group, 24 of the participants invited attempted to complete the survey, Of them, 3 failed to complete the entire survey and 7 failed to requalify in the prescreen validation. For the Control group, 38 re-qualified in the prescreen re-validation, completed the entire survey, and were not flagged as bots. Only 14 did so for the Childhood Trauma Group. 1 participant from both the Control and Childhood Trauma Group failed the attention checks and were also excluded, leaving a total of 37 and 13 participants for the Control and Childhood Trauma group respectively. Due to the low number of participants, the Childhood Trauma Group was removed from the main analysis. Tentative

analyses including the Childhood Trauma Group were run and included in Appendix C but must be treated with caution considering the small sample sizes of the group. Consequently, a final total of 72 participants were assessed for the official study, 35 in the CPTSD group and 37 in the Control group.

Power Analysis

In light of the small participant pool, a power analysis was conducted to assess the power of the study. For the original experimental design with three groups, power analysis for a repeat measure, within-between ANOVA showed that a total sample size of 120 participants would have been required to achieve a statistical power of 0.8 and a moderate effect size of .20 at a significance level of 0.05. In general, an effect size of 0.10 is considered small, 0.25 medium, and 0.40 as large (Cohen, 1969). In this analysis, the effect size for the study was computed within G*Power using means and standard deviations, coming out to be approximately 0.2.

With two groups, 96 participants would have been required to achieve the same level of statistical power, effect size, and significance. In the current study, there were 75 participants in total. Post-hoc power analysis indicated that at a sample size of 75, the study currently has an approximate power of 0.69. The results of the analysis indicate that the study may be underpowered. It is currently only powered enough to detect large to moderate effects. Some small effects may not have been detected.

Demographics

The sample population was predominately White at 76.4% (n = 55). 11.1% (n = 8) identified with multiple races, 8.3% (n = 6) identified as Asian, 2.8% (n = 2) as American Indian or Alaskan Native, and 1.4% (n = 1) identified as Other. Majority of participants identified as

female at 63.9% (n = 46) with 31.9% (n = 23) identifying as Male. 2.8% (n = 2) and 1.4% (n = 1) identified as Gender Non-binary and Other respectively. Most of the participants had some level of college-education, with only 8.3% (n = 6) reporting that their highest level of education was high school or equivalent. 30.6% (n = 22) reporting receiving at least some college education with 5.6% (n = 4) reporting an associate degree, 37.5% (n = 27) a bachelors, 15.3% (n = 11) a masters, 1.4% (n = 1) a doctoral degree, and 1.4% (n = 1) a professional degree. 68.1% (n = 49) reported that they had never married with only 26.4% (n = 19) reporting being currently married. 4.2% (n =3) were divorced and 1.4% (n = 1) were widowed.

23.6% (n = 17) made between $25,000-$49,000 a year, followed by 19.4% (n = 14) making between $50,000-$74,999 and $75,000-$99,999. Only 8.3% (n = 6) made between $100,000-$200,000. 25% (n = 18) reported making less than $25,000 a year and 4.2% (n = 3) reported making more than $200,000. 38.9% (n = 28) of participants were between the age of 25 and 34, with 20.8% (n = 15) being between 18 and 24. 18.1% (n = 13) were between 35 and 44, 15.3% (n = 11) between 45 and 54, 2.8% (n = 2) between 55 and 64, and 4.2% (n = 3) between 64 and 74.

Fisher's Exact Test was conducted to assess for differences in demographics between CPTSD and Control groups (see Table 2). There were no significant differences detected for race ($p = .154$), highest level of education attained ($p = .44$), gender ($p = .235$), relationship status ($p = .395$), and age ($p = .102$). There was, however, a significant difference in financial status ($p < .001$) between the two groups, with the CPTSD typically occupying lower income brackets than the Control group.

Table 2

Participant Demographics

	CPTSD (n = 35, 51.4%)		Control (n = 37, 48.6%)		Overall (n = 72)		*p* value
	n	%	n	%	n	%	
Race							.154
White	26	74.3	29	78.4	55	76.4	
Asian	1	2.9	5	13.5	6	8.3	
American Indian or Alaskan Native	2	5.7	-	-	2	2.8	
Multiple Selected	5	14.3	3	8.1	8	11.1	
Other	1	2.9	-	-	1	1.4	
Education							.44
High school graduate or equivalent	4	11.4	2	5.4	6	8.3	
Some college but not degree	13	37.1	9	24.3	22	30.6	
Associate degree	3	8.6	1	2.7	4	5.6	
Bachelor's degree	11	31.4	16	43.2	27	37.5	
Master's degree	4	11.4	7	18.9	11	15.3	
Doctoral degree	-	-	1	2.7	1	1.4	
Professional degree	-	-	1	2.7	1	1.4	
Gender							.235
Female	23	65.7	23	62.2	46	63.9	
Male	9	25.7	14	37.8	23	31.9	
Gender Nonbinary	2	5.7	-	-	2	2.8	
Other	1	2.9	-	-	1	1.4	
Relationship Status							.395
Never Married	27	77.1	22	59.5	49	68.1	
Married	7	20	12	32.4	19	26.4	
Divorced	1	2.9	2	5.4	3	4.2	
Widowed	-	-	1	2.7	1	1.4	
Financial Status							<.001
Less than $25,000	16	45.7	2	5.4	18	25	
$25,000 - $49,999	9	25.7	8	21.6	17	23.6	
$50,000 - $74,999	6	17.1	8	21.6	14	19.4	
$75,000 - $99,999	4	11.4	10	27	14	19.4	
$100,000 - $200,000	-	-	6	16.2	6	8.3	
More than $200,000	-	-	3	8.1	3	4.2	
Age							.102
18 – 24	11	31.4	4	10.8	15	20.8	
25 – 34	13	37.1	15	40.5	28	38.9	
35 – 44	7	20	6	16.2	13	18.1	
45 – 54	3	8.6	8	21.6	11	15.3	
55 – 64	1	2.9	1	2.7	2	2.8	
65 – 74	-	-	3	8.1	3	4.2	

Prescreen Surveys

The two groups had significantly different total scores for the CTQ and ACES. The median ACES score for the overall sample was 2 and the median CTQ score was 46. Median scores for the CTQ subscales were in the "None" grade score range, except for Emotional Neglect, which was in the "Low" grade score range at a median score of 12 (see Table 3). Mann-Whitney U tests found that the CPTSD group (Mdn = 5) had higher ACES scores compared to the Control group (Mdn = 0), $U = -6.83$, $p < .001$, $r = .805$. The same was true for CTQ scores, with the CPTSD group (Mdn = 68) having higher scores than the Control group (Mdn = 40), $U = -6.68$, $p < .001$, $r = .787$. Mann Whitney U tests were also conducted comparing CPTSD and Control scores on the individual subscales of the CTQ-SF. For the emotional abuse, emotional neglect, physical neglect, and sexual abuse subscales, the CPTSD group generally scored higher on the CTQ-subscales than the Control group (see Table 2).

However, for the CTQ subscale for Denial, the Control group (Mdn = 0) was generally found to be more likely to minimize or dismiss traumatic experiences than the CPTSD group (Mdn = 0), $U = 2.24$, $p = .031$, $r = .264$ (see Figure 3). MacDonald et al. (2016) has previously advocated in favor of removing participants with positive scores (above 0) for the denial subscale but has also noted the effect of minimization/denial is most apparent in large sample sizes. They have also reported that there is no simple cutoff score to determine valid vs invalid reporting. Participants with positive denial subscale scores were thus retained for the main analysis. The main analysis was replicated with participants removed, one where only participants who scored the maximum denial subscale score (3) were removed (see Appendix D) and one where all participants with positive denial subscale scores were removed (see Appendix E) to ensure that results were largely unaffected.

Table 3

Pre-Screen Trauma Surveys

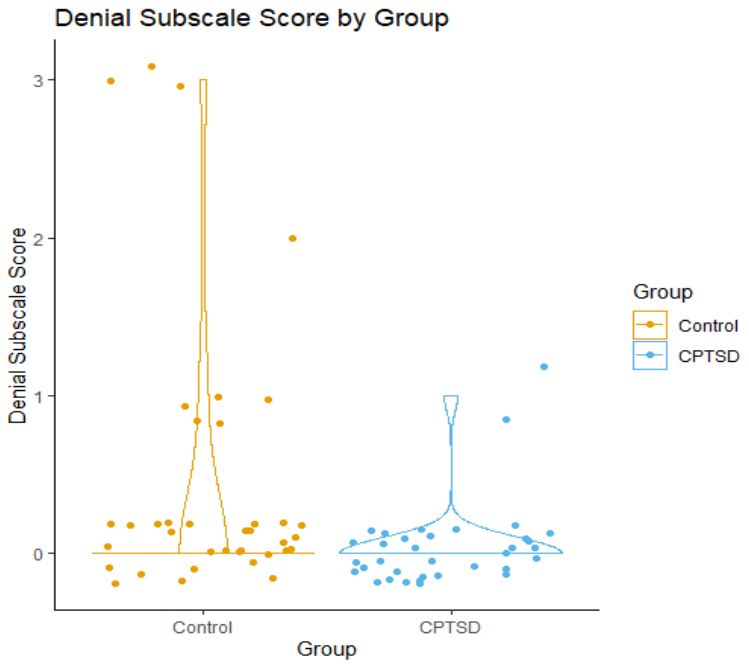
	CPTSD (n = 35, 51.4%)	Control (n = 37, 48.6%)	Overall (n = 72)	*U*	*p* value	*r*
ACEs	5	0	2	-6.83	<.001	.805
CTQ	68	40	46	-6.68	<.001	.787
Emotional Abuse	18	6	8	-6.83	<.001	.805
Emotional Neglect	18	8	12	-6.30	<.001	.742
Physical Abuse	8	5	6	-5.35	<.001	.630
Physical Neglect	11	5	7	-6.54	<.001	.770
Sexual Abuse	8	5	5	-5.45	<.001	.643
Denial[a]	0	0	0	2.24	.031	.264

Note. Values reflect medians due to non-parametric distribution. *p* values were derived from Mann Whitney U Tests that were conducted to compare CPTSD and Control groups. Wilcoxon effect sizes (r) in these cases exhibit large effects ($r \geq .5$) except for the Denial subscale, which exhibited a small effect ($.1 < r < .3$)

[a] Medians for subscale were identical but distribution differed. See Figure 3 for the direction of significance.

Figure 3

Distribution of Denial Subscale



Note. The CTQ Denial Subscale measures underreporting of trauma, with increasing minimization as the score increases, with a max score of 3 and a minimum score of 0. Data points were jittered to prevent overlap—any points below 0 on the graph indicate a denial subscale score of 0.

Impression

A three-way mixed ANOVA with a 2x2x3 design was run with Group (CPTSD vs. Control) as the between-subject variable and Agent (Bad vs. Good) and Phase (Formation vs. Violation vs. Reparation) as the within-subject variables for impression rating (see Table 4). Tests ran to ensure that the assumptions for ANOVA were met found that the data was normally distributed, as evaluated using the Shapiro-Wilk's test for normality ($p > .05$), and that there was homogeneity of variance, which was tested for using Levene's test ($p > .05$). Mauchly's test for Sphericity indicated violations of assumptions for Phase ($W = 0.774$, $p < .001$) and Group x Phase ($W = 0.774$, $< .001$). Degrees of freedom were then subsequently corrected for using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.815$).

Table 4

ANOVA Results for Impression

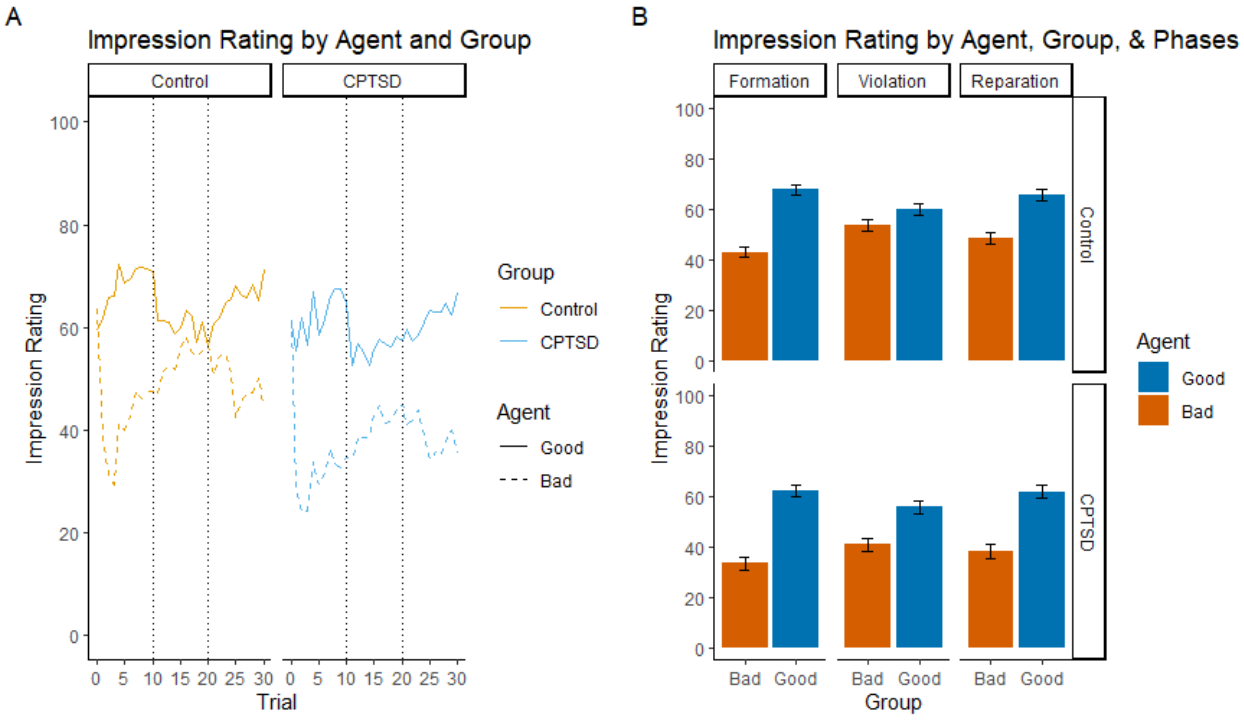
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 70	3.07	.042	.084	.
Phase	1.63, 114.2	2.52	.035	.096	.
Group : Phase	1.63, 114.2	0.416	.006	.619	
Agent	1, 70	92.1	.568	<.001	***
Group : Agent	1, 70	2.51	.035	.118	
Phase : Agent	1.95, 136.5	40.3	.365	<.001	***
Group : Phase : Agent	1.95, 136.5	0.77	.011	.463	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes ($.01 < \eta_p^2 < .06$) for Group, Phase, Group x Agent, and Group x Phase x Agent. Large effects ($\eta_p^2 > .14$) were present for Agent and Phase x Agent.

The ANOVA found significant main effects for Agent, $F(1, 70) = 92.1$, $p < .001$, $\eta_p^2 = .568$, along with a significant Phase x Agent interaction, $F(1.95, 136.5) = 40.3$, $p < .001$, $\eta_p^2 = .365$. Figure 4A reflects impression ratings by Group and Agent over the Trials while Figure 4B reflects the same by Phase. Graphs indicated less favorable impression ratings for the bad agent than the good agent. Similarly, the CPTSD group also appeared to rate the agents slightly lower than the Control group. In the Violation phase, changes in impression rating for the bad agent appear to be less drastic for the CPTSD group vs the Control group.

Figure 4

Impression Ratings by Agent and Group throughout Task

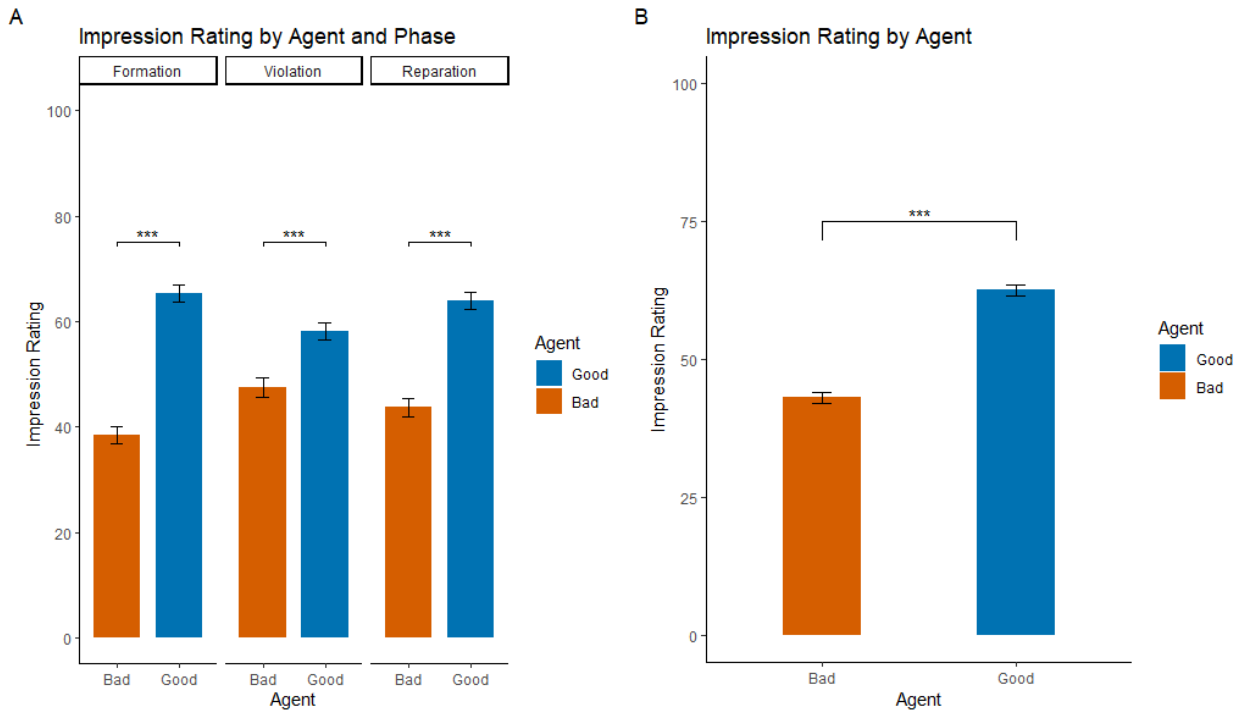


Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Post-hoc analysis suggested that bad agents (M = 43.1, 95% CI [42, 44.1]) scored significantly lower in overall impression ratings compared to Good Agents (M = 62.5, 95% CI [61.6, 63.5]), $t(70) = -9.59$, $p < .001$ (see Figure 5B). This trend held true within each phase (see Figure 5A). Bad agents (M = 38.5, 95% CI [36.8, 40.2]) scored significantly lower than good agents (M = 65.3, 95% CI [63.8, 66.8]) in the Formation phase, $t(70) = -12.4$, $p < .001$. In the Violation phase, good agents (M = 58.1, 95% CI [56.4, 59.8]) once again scored higher compared to the bad agents (M = 47.5, 95% CI [45.7, 49.4]), $t(70) = -4.51$, $p < .001$. This trend was replicated in the Reparation phase, with the bad agents (M = 43.7, 95% CI [41.9, 45.5]) scoring lower than the good agents (M = 63.9, 95% CI [62.2, 65.6]), $t(70) = -8.97$, $p < .001$.

Figure 5

Significant Main Effects and Interactions for Impression Rating



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Certainty

An Aligned Rank Transformation ANOVA (ART ANOVA) was conducted for certainty rating, with Group (CPTSD vs. Control) as a between-subject variable and Agent (Bad vs. Good) and Phase (Formation vs. Violation vs. Reparation) as within-subject variables (see Table 5) as assumptions for a parametric ANOVA were violated. The ART ANOVA found a significant two-way interaction for Group x Agent, $F(1, 350) = 15.8$, $p < .001$, $\eta_p^2 = .043$.

Table 5

ART ANOVA Results for Certainty

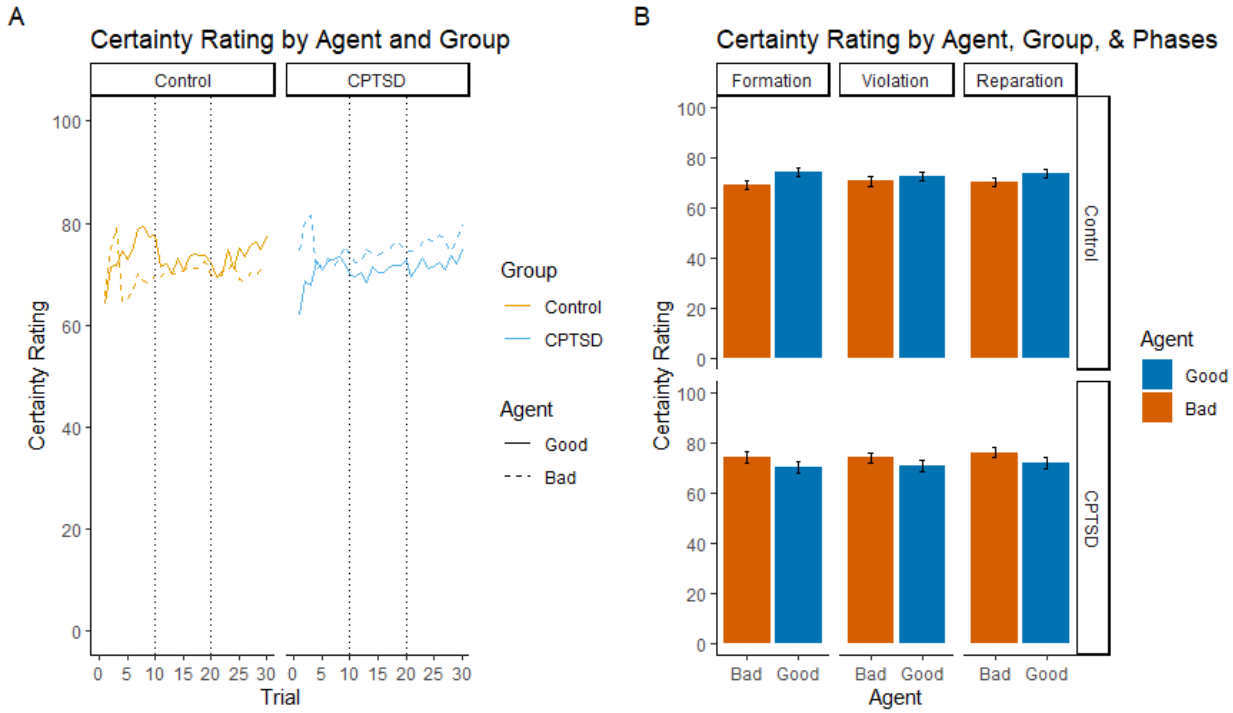
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 70	0.191	.003	.663	
Phase	2, 350	0.314	.002	.731	
Group : Phase	2, 350	0.160	.001	.852	
Agent	1, 350	0.104	<.001	.748	
Group : Agent	1, 350	15.8	.043	<.001	***
Phase : Agent	2, 350	0.199	.001	.819	
Group : Phase : Agent	2, 350	0.719	.004	.488	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) indicated a small effect size ($.01 < \eta_p^2 < .06$) for Group x Agent.

Figure 6A reflects impression ratings by Group and Agent over the Trials while Figure 6B reflects the same by Phase. Figures indicate that the CPTSD group was more certain about the bad agent than the good agent, with the reversal being true for the Control group.

Figure 6

Certainty Ratings by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Corresponding post-hoc analysis found that both CPTSD and Control groups reported significant differences in certainty ratings for good and bad Agents (see Figure 7). The Control group generally reported lower certainty ratings for the bad agent (M = 70.1, 95% CI[69.0, 71.1]) vs the good agent (M = 73.5, 95% CI [72.5, 74.5]), $t(350) = -2.70$, p = .044. However, the CPTSD group generally reported higher certainty ratings for the bad agent (M = 75.0, 95% CI [73.8, 76.3]) and lower certainty ratings for the good agent (M = 71.0, 95% CI [69.7, 72.3]), $t(350) = 2.91$, p = .023.

Figure 7

Significant Main Interaction for Certainty Ratings

Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Prediction Accuracy

To assess prediction accuracy, an ART ANOVA was run with Group (CPTSD vs. Control) as a between-subject variable and Agent (Bad vs. Good) and Phase (Formation vs. Violation vs. Reparation) as within-subject variables as assumptions were not met for a parametric ANOVA (see Table 6). It observed a significant main effect for Phase, $F(2, 350) = 8.17$, $p < .001$, $\eta_p^2 = .043$, and a significant two-way interaction for Agent x Phase, $F(2, 350) =$

39.3, p <.001, η_p^2 = .183. It did not find a significant three-way interaction, only a marginally significant one, F(2, 350) = 2.39, p = .093, η_p^2 =.013.

Table 6

ART ANOVA Results for Prediction Accuracy

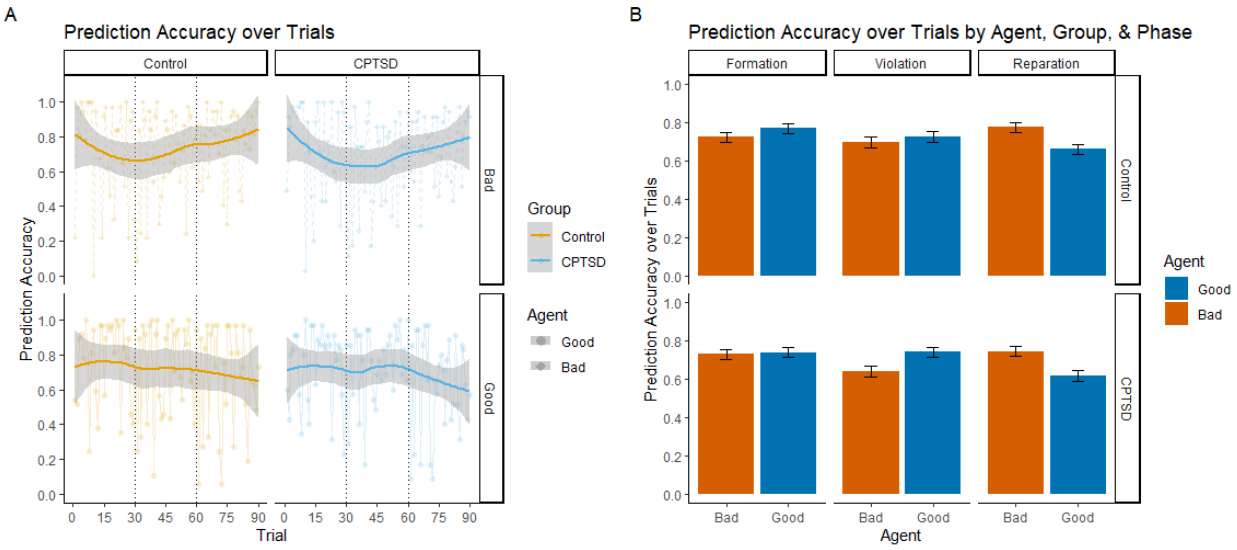
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 70	2.54	.035	0.116	
Phase	2, 350	8.17	.045	<.001	***
Group : Phase	2, 350	0.481	.003	.619	
Agent	1, 350	1.24	.003	.266	
Group : Agent	1, 350	.022	<.001	.881	
Phase : Agent	2, 350	39.3	.183	<.001	***
Group : Phase : Agent	2, 350	2.39	.013	.093	.

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes (.01 < η_p^2 < .06) for Group, Phase, and Group x Phase x Agent. Large effects (η_p^2 <.14) were present for Phase x Agent.

Figure 8A reflects impression ratings by Group and Agent over the Trials while Figure 8B reflects the same by Phase. Both the CPTSD and Control group appeared to follow similar patterns of prediction accuracy for the good and bad agents.

Figure 8

Prediction Accuracy by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Average Prediction Accuracy per Trial was mapped in the background and a local polynomial regression (solid line) was fitted over it. Error bands represent 95% CI. (B) Error bars represent 95% CI.

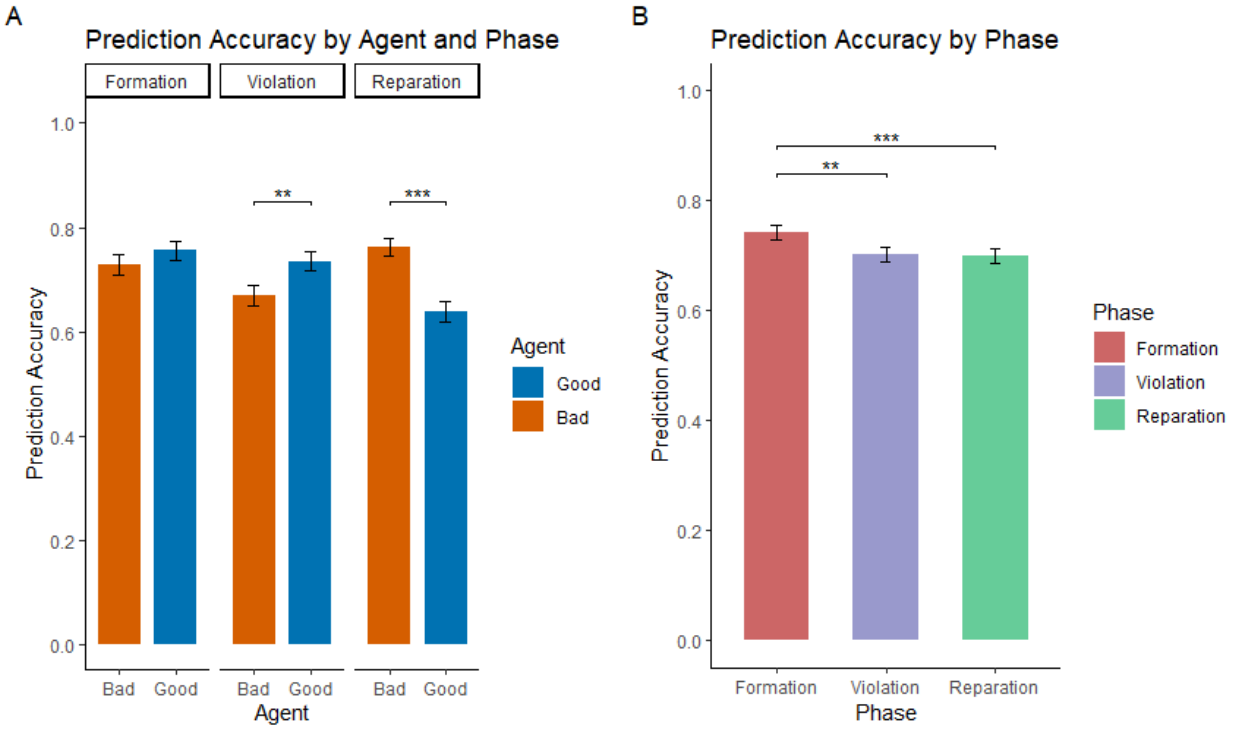
Post-hoc analysis for the two-way interaction for Agent x Phase found significant differences in prediction accuracy between bad and good agents during the Violation and Reparation phases (see Figure 9A). In the Violation phase, bad agents (M =0.67, 95% CI [0.65, 0.69]) performed worse than good agents (M = 0.735, 95% CI [0.716, 0.754]), $t(350) = -3.77$, $p = .003$. In contrast, in the Reparation phase, bad agents (M = 0.762, 95% CI [0.744, 0.78]) performed better than good agents (M = 0.639, 95% CI [0.619, 0.659]), $t(350) = 7.96$, $p < .001$. There was, however, no significant difference between good (M = 0.756, 95% CI [0.738, 0.774]) and bad agents (M = 0.728, 95% CI [0.709, 0.747]) during the Formation phase, $t(350) = -2.11$, $p = .534$.

For Phase, post-hoc analysis found significant differences between the Formation and Violation phase, $t(350) = 3.24$, $p = .004$, as well as between the Formation and Reparation phase,

t(350) = 3.71, p <.001. It did not find a significant difference between the Violation and Reparation phase, *t*(350) = 0.472, p = 1.00 (see Figure 9B). In general, it appears that prediction accuracy was lower for the Violation phase (M = 0.702, 95% CI [0.688, 0.716]) compared to the Formation phase (M = 0.742, 95% CI [0.729, 0.755]). The Formation phase had greater prediction accuracy than the Reparation phase (M =0.7, 95% CI [0.686, 0.714]).

Figure 9

Significant Main Effects and Interactions for Prediction Accuracy



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Discussion

The study attempted to compare moral inferencing patterns between participants with CPTSD symptoms and healthy controls. Results found a significant Group x Agent interaction for certainty ratings. The analysis also detected significant main effects for Agent and a

significant Agent x Phase interaction for impression rating. Additionally, it also detected a significant main effect for Phase and a significant Agent x Phase interaction for prediction accuracy.

Group Statistics

Within our study, 13.88 % (n = 89) of the participants initially screened met criteria for CPTSD. Of those who later participated in the main survey, however, only 42 (60%) met criteria a second time. Due to the relative newness of the disorder, there have not been reliable prevalence rates established for CPTSD. Cloitre et al. (2018) found that 12.9% of participants met criteria for CPTSD in their online community sample when validating the ITQ. Amongst clinical samples, however, Cloitre et al. (2018) found rates as high as 61.1%. An alternative population study for CPTSD that looked at noninstitutionalized U.S. adults, on the other hand, reported prevalence rates of about 3.8% for CPTSD (Cloitre et al., 2019). The prevalence rate of CPTSD found within the study overall matches closely to the rates detected by Cloitre et al. (2018) in online community samples.

In general, there were no significant demographic differences found between the two groups outside of their financial status. However, the difference was not entirely unexpected. CPTSD, and mental illness in general, have often been correlated with poorer financial outcomes (Daniunaite et al., 2021). Unexpectedly, there was not a significant difference in gender between the CPTSD and Control group for the study. Cloitre et al. (2019) had previously reported higher rates of CPTSD among female participants versus male participants. However, considering the sample size of the study, the discrepancy may simply be a result of insufficient power.

The differences in ACES scores, CTQ-SF scores, and CTQ subscale scores between the Control and CPTSD suggest that the criteria used to sort participants for each group were successful. In general, the CPTSD group reported more extensive trauma histories than the Control group. The Control group, however, was more likely to minimize or deny their traumatic experiences compared to Control group participants. Interestingly, both groups scored the highest on the Emotional Abuse and Emotional Neglect subscales, though for both subscales, the CPTSD group scored significantly higher than the Control group. Considering the interpersonal deficits and emotional dysregulation that characterizes CPTSD, the scoring matches current conceptualizations of CPTSD.

Impression

There were significant main effects reported for Agent as well as a significant Phase x Agent interaction. Participants tended to rate the bad agent lower than they did the good agent, which suggests that the experimental manipulations were successful. Similarly, there were significant differences between impression ratings for each agent during each phase. As agents were manipulated to have shifting moral preferences for each phase, the results suggest that manipulations were executed correctly. Notably, for both the good and bad agents in the Violation phase, while there were, on average, shifts into the opposing direction as expected by the manipulation, impression ratings never reached the extremes of the other agent. For instance, while impression ratings improved for the bad agent during the Violation phase when they exhibited moral preferences more aligned with that of the good agent, impression ratings did not improve to the extent that they received the same scores as the good agent. Current results for impression ratings suggest that experimental manipulation was successful and that the effects observed are accurate.

Hypothesis 1 predicted that differences in impression ratings for good and bad agents would be less extreme for individuals with CPTSD. In short, it predicted differences in impression rating for agents by group. A significant Group x Agent interaction, however, was not observed, which suggests that there were no differences in impression ratings between good and bad agents by group. As such Hypothesis 1 can be refuted based on the results of the current study. Notably, when participants who scored a 3 on the denial subscale were removed from the study, a significant Group x Agent interaction was observed (see Appendix D). This effect was not observed when all participants with positive denial subscale scores were removed, however, suggesting that this is not robust. Hypothesis 3, on the other hand, predicted that the CPTSD group would exhibit minimal differences in impression rating across phases. A significant Group x Phase interaction for impression ratings was not present within the current study, leading to the rejection of Hypothesis 3. There were no differences between the groups regarding impression rating across phases.

Based on the current data and analysis, no differences between impression ratings for the good and bad agents by group were observed. Nor were there differences between impression ratings across phases by group. However, the study is notably underpowered and, as such, some effects may not be fully detected. Marginally significant differences were detected for both Group and Phase in the mixed three-way ANOVA. Similarly, a significant Group x Agent interaction was detected in some supplementary analyses (see Appendix D). If the study had greater power, other trends may have been detected.

Siegel et al. (2020)'s study matched the current study's results. They reported a significant main effect for Agent but not a significant main effect for Group. Nor did they report a significant Group x Agent interaction. The results of these two studies suggest that impression

ratings of CPTSD and BPD populations do not differ drastically from that of healthy controls. As such, based on the data available, it could be speculated that both disorders do not lead to abnormalities in impression formation.

Certainty

Hypothesis 2 predicted that there would be differences in certainty ratings for the agents by group. Specifically, it predicted that participants in the CPTSD group would have little to no differences in certainty ratings between the good and bad agents relative to the Control group. Based on the data of the study, the hypothesis was partially correct. There was a significant Group x Agent interaction, which suggested that there were differences in certainty ratings between the agents for each group.

Post-hoc analysis, however, did not find that there were insignificant differences in certainty ratings for the agents in the CPTSD group compared to significant differences in certainty ratings for the agents in the Control group. For the CPTSD group, the post-hoc analysis found that participants were generally more certain about the bad agent and significantly less certain about the good agent (see Figure 7). In contrast, participants in the Control group were more certain about their impression ratings for the good agent and less certain about their impression ratings for the bad agent. Contrary to the hypothesis, there were differences in certainty ratings for the agents for the CPTSD group as well as the Control group.

Siegel et al. (2020) also detected a Group x Agent interaction when comparing untreated BPD patients with healthy controls. Unlike the results of this study, however, Siegel et al. (2020) did not detect significant differences in certainty ratings between good and bad agents within untreated BPD patients. Instead, Siegel et al. (2020) found that untreated BPD participants were

significantly more certain about the bad Agent when compared to the Control group but were equally certain about the good Agent

While the study is underpowered and no strong claim can be made, the differences present suggest that while individuals with CPTSD and healthy controls are minimal, the two groups differ in which agent they are more certain about. CPTSD groups are generally more confident about their impressions of bad agents while healthy controls are more confident about good agents. Yet, ultimately, certainty ratings between good agents do not differ between groups and neither do their ratings for bad agents. In practicality, the results suggest that CPTSD participants don't differ from Control groups overall outside of this preferential leaning whilst Siegel et al. (2020)'s results point to differences between BPD and healthy controls regarding their certainty for the bad agent.

Interestingly enough, the resulting pattern with certainty ratings detected within this current study was more aligned with Siegel et al. (2019) than Siegel et al. (2020). Siegel et al. (2019) examined moral inferencing patterns in connection with exposure to violence (ETV), as measured via a scale, in incarcerated males. They found that the higher the participants were on the ETV scale, the more certain participants were about their impressions of the bad agent and the more uncertain they were about the good agent. Notably, the pattern observed among CPTSD participants matched that of the participants in Siegel et al. (2019)'s study that reported an ETV scale score of 13 (the maximum score possible on the scale). This is not entirely unexpected, considering rates of trauma amongst the CPTSD group. Many participants within the CPTSD group were likely exposed to violence in their everyday lives.

The ART ANOVA for the certainty ratings also did not find any other significant main effects or interactions. There were no significant differences in certainty rating across Phase,

Group, or Agent. This is not in line with the previous literature, as Siegel et al. (2018) had previously reported significant differences in certainty ratings between agents. Previously, the literature had pointed to lower certainty for the bad agent and higher certainty for the good agent. In this study, no such significant main effect was detected. Contradictorily, certainty ratings seem relatively similar for both, suggesting that either perhaps experimental manipulation was not explicit enough or that certainty ratings remained largely unaffected, or minimally affected, by the current experimental manipulations. Considering that there were frequent moral preferences shifts throughout the experiment, it could be that participants remained relatively uncertain throughout and never had opportunities to form concrete impressions that they were certain of.

Prediction Accuracy

Hypothesis 4 further predicted that CPTSD groups may have lower overall prediction accuracy than others. There was, however, no significant main effect detected for Group. As such, this hypothesis can be rejected as there were no significant differences in prediction accuracy between groups. The ART ANOVA did, however, detect a marginally significant interaction between Group x Agent x Phase. In supplementary analyses, when participants with positive denial subscale scores were removed, this effect was significant (see Appendix E). The same effect was observed when only participants with a denial subscale score of 3 were removed (Appendix D). As the study is underpowered, with more data, a clearer effect may be detected with more participants.

The ANOVA, however, did detect a significant main effect for Phase as well as an Agent x Phase interaction. Significant differences between phases were not altogether surprising as

each phase involved an experimentally manipulated abrupt change in an agent's moral preferences. The post-hoc analysis found significant differences between the Formation and Violation phase as well as the Formation and Reparation phase but not between Violation and Reparation (see Figure 9B). The differences can be attributed to the changes in moral preferences, which may violate the participant's expectations about the agent's behavior during those phases.

The Agent x Phase interaction suggested that there were differences in prediction accuracy by agent in each phase. There were no differences in prediction accuracy for each agent in the Formation phase, but differences were detected between the agents for the Violation and Reparation phase (see Figure 9A). Prediction accuracy was significantly lower for the bad agent in the Violation phase but higher for the good agent in the Reparation phase. The dynamic was reversed in the Reparation phase, with greater accuracy for the bad agent and significantly lower accuracy for the good agent. Notably, in both cases, prediction accuracy was higher for agents that were exhibiting behavior that was aligned with a bad agent and lower for agents exhibiting behavior that was aligned with a good agent.

This pattern matches that of the literature. Past research has long since established that negative information is weighed more than positive information and that threat-related cues typically increase attention and information processing (Siegel et al., 2018). In one of the studies conducted in Siegel et al. (2018), they assessed whether new negative impression information would mediate faster impression updating using a modified version of the Moral Inferencing Task. To do so, they recruited 408 healthy controls and had them learn randomly about a good or bad agent through 36 trials. Within the last six trials, they altered the moral preferences of the agents by shifting their κ by 0.2 in either direction, therefore making the agents either more

harm-averse or less harm averse. The study found that not only were participants faster to update their impressions of bad agents, but they were also generally faster to do so when the experimental manipulation led agents to be less harm-averse. In the current study's case, the patterns in prediction accuracy are potentially attributable to the fact that participants are faster to update negative information and are therefore more accurate in their assessment of agents exhibiting negative behavior. However, as the study did not evaluate impression or belief updating, no concrete explanation can be made for the patterns observed.

Interpersonal Implications for CPTSD

The study largely found group differences in certainty ratings for each agent. The results generally suggest that individuals with CPTSD are quicker to be confident about negative impressions they hold toward people but are conversely less certain about positive information. This matches current conceptualizations of CPTSD individuals, who typically view others as dangerous or unpredictable (Nieuwenhove & Meganck, 2019). Bad agents confirm this preconception for CPTSD patients while good agents do not. Individuals with CPTSD may remain uncertain about good agents and be attentively biased toward negative information. They may, in effect, be actively looking for information to confirm their negative prior expectations and to validate their preferences for avoiding relationships. Notably, the patterns observed within CPTSD populations are more consistent with those exposed to violence versus those diagnosed with BPD.

Outside of relationship avoidance, interpersonal difficulties among CPTSD patients are also characterized by maladaptive relationships. In this case, this study finds promising results in the sense that there were no group differences in prediction accuracy for each agent per phase.

There is no evidence to suggest that the CPTSD group is not just as equally accurate in their assessment of good and bad agents through the phases as the Control group. As such, the CPTSD group, like healthy controls, may have the informational processing capacity to detect changes in moral preferences and character. The capability of detecting changes in the moral character of others can be essential for assessing the maladaptively of interpersonal relationships.

While further research would need to be conducted to confirm this, the results are promising in the sense that they suggest that interpersonal difficulties in CPTSD patients are not fueled by the inability to detect changing moral characters of others. The study also further suggests that CPTSD, like BPD, does not impair one's ability to form impressions of the moral character of others as no significant main effect for group was detected for impression rating. Based on the results of this study, it appears that while deficits in moral inferencing may potentially explain tendencies towards relationship avoidance, such deficits are insufficient to explain why CPTSD patients tend to engage with maladaptive relationships.

Limitations

This study was foremostly limited by the small sample size and, subsequently, the low statistical power. Due to the strict cutoffs used to evaluate participants for experimental groups, few participants successfully qualified for the study. As such, each arm of the study was therefore insufficiently powered, and it is unclear how credible the results of the study current are. With a larger sample size, the effects evident in the current study may be further amplified and a clearer conclusion may be drawn. The small sample size can partially be attributable to the attrition that occurred within the waiting period between the initial pre-screen and the invitation sent for the main survey. In the future, minimizing the wait time may improve participant

responsiveness and lead to larger sample sizes. In this case, as criteria for Control and Trauma groups were partially derived from the Prescreen data, the wait time was inevitable, but could no doubt have been shortened. In the future, it would also be advisable to recruit within clinical populations and to severely increase the scope of the pre-screening survey to capture a larger sample pool.

The external validity of this experiment is also under question due to the demographic spread of the participants. Current census data for the US population reports that 76.3% of the population identifies as White, 18.5% as Hispanic or Latino, 13.4% as Black or African American, 5.9% Asian, 1.3% as American Indian or Alaskan Native, 0.2% as Native Hawaiian or Pacific Islander, and 2.8% as identifying with two or more races (U.S. Census Bureau QuickFacts, n.d.). The current participant pool matches percentage levels for White populations but does not have sufficient representation for other racial/ethnic groups in the U.S. There is noticeably a dearth of Black, Hispanic, and/or Latino participants within the study that make it clear that the study itself may not be representative of U.S. populations. As such, it is debatable whether the implications of the study can apply broadly to the U.S. Similarly, because of the low numbers of Black, Indigenous, and People of Color (BIPOC) individuals within this study, the results are not generalizable to such minority groups.

The online nature of the study allowed for ease of distribution and provided access to a greater range of participants but also introduced confounding variables into the study. As each participant completed the Moral Inferencing Task virtually, it was impossible to monitor them to ensure attentiveness. Participants may have been distracted and unfocused, therefore perhaps prone to careless/insufficient effort (C/IE) responding. Considering that the study in question was 40 minutes in duration and repetitive in nature, it is likely that there was some of the data

may have been compromised as a result. Following recommendations by Zorowitz et al. (2021), the study included four infrequency attention check questions to evaluate for such instances. Two infrequency attention check questions were included in the self-report section of the survey and two within the Moral Inferencing Task itself. These attention check questions were added to the study halfway through data collection. As a result, 21 participants in the CPTSD group did not encounter any attention checks. However, very few participants were found to have actually failed the attention checks in general, suggesting that most participants were attentive and that the data is largely reliable.

Similarly, because the study was conducted online, self-report measures were used to assess eligibility for the groups. It is unclear how accurate such measures are in assessing the symptomology. Some participants that were screened and met criteria for CPTSD and/or other groups did not revalidate their symptomology in the main survey and were thus excluded. The number of participants that failed to revalidate in the main survey, however, is alarming and suggests that the self-report measures utilized to determine eligibility may have low test-retest reliability. For the CPTSD group, of the 70 who re-participated, 28 failed. For the Control group, similarly, 15 out of the 59 that attempted the survey failed to requalify while 7 of the 24 for the Childhood Trauma group also failed to revalidate. Past research has suggested that online self-report measures generally correlate well with traditional paper assessment but has also noted a retesting effect, where participants generally report lower scores when retested later on (Luce et al., 2007). This effect may explain the high rates of disqualification in the main survey. The participants included in the analysis, however, met the eligibility criteria twice, which grants some reassurance, but the validity of the self-report questionnaires remains debatable.

Future Directions

Due to the low sample size and power for this study, in the future, the study would need to be replicated with a larger sample size to ensure that replication of results would be possible. In the current study, only impression ratings, certainty ratings, and prediction accuracy were examined but it may be beneficial to model belief updating in future studies to better understand how moral inferencing works within such groups. In this case, prediction accuracy was used to make inferences about belief updating patterns in CPTSD patients but modeling learning rates may provide clearer insight into how CPTSD patients evaluate and update beliefs about others,

The results of the study currently suggest that there are minimal deficits in impression formation (outside of some abnormal patterns in impression certainty). There is, further, no evidence suggesting that CPTSD populations are less able to alter impressions as needed with new information. As such, current preliminary evidence suggests that there are no deficits within moral inferencing itself and alternative avenues may be worth evaluating. Specifically, it may be interesting to evaluate if CPTSD individuals' behaviors match cognitive understandings of impressions. As the Moral Inferencing Task measures cognitive understandings of people's moral character, the study can effectively evaluate whether participants have deficits in perceiving the moral character of others. However, cognitive understanding does not necessarily mean that a participant's behavior will match such understandings accordingly. An individual may cognitively understand that a stranger is dangerous yet behaviorally, they may still engage in interaction.

Future studies can evaluate whether there is a discrepancy present. Previous studies conducted by Siegel et al. (2018) have tested whether participants' moral impressions of agents

affected their social behavior. Siegel et al. (2018) tested healthy controls in that case using a trust game and found that participants generally trusted more money to the good agent versus the bad agent. It is unclear whether the same pattern would emerge for CPTSD participants. Abramov et al. (2020)'s used multi-round trust games with participants with BPD symptomology in a similar Formation-Violation-Reparation paradigm except with trust. In that study, Abramov et al. (2020) observed abnormal patterns amongst High-BPD symptom groups, where high BPD participants tend to invest more money after trust violations and, counterintuitively, invest less money after trust reparation. A future study testing the relation between moral inferencing and the social behavior of CPTSD participants may shed more light on the mechanisms behind impaired interpersonal relations in CPTSD.

References

Abramov, G., Miellet, S., Kautz, J., Grenyer, B. F. S., & Deane, F. P. (2020). The paradoxical decline and growth of trust as a function of borderline personality disorder trait count: Using discontinuous growth modelling to examine trust dynamics in response to violation and repair. *PLOS ONE*, *15*(7), e0236170. https://doi.org/10.1371/journal.pone.0236170

Alboukadel, K. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [R package version 0.7.0]. https://CRAN.R-project.org/package=rstatix

Bernstein, D. P., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report manual*. The Psychological Corporation.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Handelsman, L., Medrano, M., Desmond, D., & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, *27*(2), 169–190. https://doi.org/10.1016/S0145-2134(02)00541-0

Bertó, C., Ferrin, M., Barberá, M., Livianos, L., Rojo, L., & García-Blanco, A. (2017). Abnormal emotional processing in maltreated children diagnosed of Complex Posttraumatic Stress Disorder. *Child Abuse & Neglect*, *73*, 42–50. https://doi.org/10.1016/j.chiabu.2017.09.020

Cloitre, M., Garvert, D. W., Brewin, C. R., Bryant, R. A., & Maercker, A. (2013). Evidence for proposed ICD-11 PTSD and complex PTSD: A latent profile analysis. *European Journal of Psychotraumatology*, *4*(1), 20706. https://doi.org/10.3402/ejpt.v4i0.20706

Cloitre, M., Garvert, D. W., Weiss, B., Carlson, E. B., & Bryant, R. A. (2014). Distinguishing PTSD, Complex PTSD, and Borderline Personality Disorder: A latent class analysis. *European Journal of Psychotraumatology*, *5*(1), 25097. https://doi.org/10.3402/ejpt.v5.25097

Cloitre, M., Hyland, P., Bisson, J. I., Brewin, C. R., Roberts, N. P., Karatzias, T., & Shevlin, M. (2019). ICD-11 Posttraumatic Stress Disorder and Complex Posttraumatic Stress Disorder in the United States: A Population-Based Study. *Journal of Traumatic Stress*, *32*(6), 833–842. https://doi.org/10.1002/jts.22454

Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, *138*(6), 536–546. https://doi.org/10.1111/acps.12956

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* Academic Press.

Daniunaite, I., Cloitre, M., Karatzias, T., Shevlin, M., Thoresen, S., Zelviene, P., & Kazlauskas, E. (2021). PTSD and complex PTSD in adolescence: Discriminating factors in a population-based cross-sectional study. *European Journal of Psychotraumatology*, *12*(1), 1890937. https://doi.org/10.1080/20008198.2021.1890937

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Ford, J. D., & Courtois, C. A. (2014). Complex PTSD, affect dysregulation, and borderline personality disorder. *Borderline Personality Disorder and Emotion Dysregulation*, *1*(1), 9. https://doi.org/10.1186/2051-6673-1-9

Herman, J. L. (2015). *Trauma and recovery: The aftermath of violence—From domestic abuse to political terror*.

Hothorn, T., Hornik, K., van de Wiel, M., & Zeileis, A. (2006). *A Lego system for conditional inference.* (60 (3)) [Computer software]. https://doi.org/10.1198/000313006X118430

Hyland, P., Karatzias, T., Shevlin, M., & Cloitre, M. (2019). Examining the Discriminant Validity of Complex Posttraumatic Stress Disorder and Borderline Personality Disorder Symptoms: Results From a United Kingdom Population Sample. *Journal of Traumatic Stress*, *32*(6), 855–863. https://doi.org/10.1002/jts.22444

Kay, M., Elkins, L., Higgins, J., & Wobbrock, J. (2021). *ARTool: Aligned Rank Transform for Nonparametric Factorial ANOVAs* [R package version 0.11.1]. https://github.com/mjskay/ARTool>

Laddis, A. (2010). Outcome of crisis intervention for borderline personality disorder and post traumatic stress disorder: A model for modification of the mechanism of disorder in complex post traumatic syndromes. *Annals of General Psychiatry*, *9*(1), 19. https://doi.org/10.1186/1744-859X-9-19

Luce, K. H., Winzelberg, A. J., Das, S., Osborne, M. I., Bryson, S. W., & Taylor, C. B. (2007). Reliability of self-report: Paper versus online administration. *Computers in Human Behavior*, *23*(3), 1384–1389. https://doi.org/10.1016/j.chb.2004.12.008

Luyten, P., Campbell, C., & Fonagy, P. (2019). Borderline personality disorder, complex trauma, and problems with self and identity: A social-communicative approach. *Journal of Personality*, *88*(1), 88–105. https://doi.org/10.1111/jopy.12483

MacDonald, K., Thomas, M. L., Sciolla, A. F., Schneider, B., Pappas, K., Bleijenberg, G., Bohus, M., Bekh, B., Carpenter, L., Carr, A., Dannlowski, U., Dorahy, M., Fahlke, C., Finzi-Dottan, R., Karu, T., Gerdner, A., Glaesmer, H., Grabe, H. J., Heins, M., … Wingenfeld, K. (2016). Minimization of Childhood Maltreatment Is Common and Consequential: Results from a Large, Multinational Sample Using the Childhood Trauma Questionnaire. *PLOS ONE*, *11*(1), e0146058. https://doi.org/10.1371/journal.pone.0146058

Maercker, A. (2021). Development of the new CPTSD diagnosis for ICD-11. *Borderline Personality Disorder and Emotion Dysregulation*, *8*, 1–4. http://dx.doi.org/10.1186/s40479-021-00148-8

Nieuwenhove, K. V., & Meganck, R. (2019). Interpersonal Features in Complex Trauma Etiology, Consequences, and Treatment: A Literature Review. *Journal of Aggression, Maltreatment & Trauma*, *28*(8), 903–928. https://doi.org/10.1080/10926771.2017.1405316

Pearlman, L. A., & Courtois, C. A. (2005). Clinical applications of the attachment framework: Relational treatment of complex trauma. *Journal of Traumatic Stress*, *18*(5), 449–459. https://doi.org/10.1002/jts.20052

R Core Team. (2019). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org/

Russell, V. L. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means* [R package version 1.7.2.]. https://CRAN.R-project.org/package=emmeans

Siegel, J., Curwell-Parry, O., Pearce, S., Saunders, K., & Crockett, M. (2020). A Computational Phenotype of Disrupted Moral Inference in Borderline Personality Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*. https://doi.org/10.1016/j.bpsc.2020.07.013

Siegel, J., Estrada, S., Crockett, M., & Baskin-Sommers, A. (2019). Exposure to violence affects the development of moral impressions and trust behavior in incarcerated males. *Nature Communications*, *10*, 1234567890. https://doi.org/10.1038/s41467-019-09962-9

Siegel, J., Mathys, C., Rutledge, R., & Crockett, M. (2018). Beliefs About Bad People Are Volatile. *Nature Human Behaviour*, *2*. https://doi.org/10.1038/s41562-018-0425-1

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (n.d.). *afex: Analysis of Factorial Experiments* (Version 2021) [R package version 1.0-1]. https://CRAN.R-project.org/package=afex

U.S. Census Bureau QuickFacts: United States. (n.d.). Retrieved April 15, 2022, from https://www.census.gov/quickfacts/fact/table/US/PST045221

Vang, M. L., Ben-Ezra, M., Shevlin, M., & Ben-Ezra, M. (2019). Modeling Patterns of Polyvictimiztion and Their Associations with Posttraumatic Stress Disorder and Complex Posttraumatic Stress Disorder in the Israeli Population. *Journal of Traumatic Stress*, *32*(6), 843–854. https://doi.org/10.1002/jts.22455

Vasilopoulou, E., Karatzias, T., Hyland, P., Wallace, H., & Guzman, A. (2020). The Mediating Role of Early Maladaptive Schemas in the Relationship between Childhood Traumatic Events and Complex Posttraumatic Stress Disorder Symptoms in Older Adults (>64 Years). *Journal of Loss and Trauma*, *25*(2), 141–158. https://doi.org/10.1080/15325024.2019.1661598

Wade, R., Becker, B. D., Bevans, K. B., Ford, D. C., & Forrest, C. B. (2017). Development and Evaluation of a Short Adverse Childhood Experiences Measure. *American Journal of Preventive Medicine*, *52*(2), 163–172. https://doi.org/10.1016/j.amepre.2016.09.033

Wickham, H. (2016). *Elegant Graphics for Data Analysis.*

Zamir, O., & Lavee, Y. (2014). Psychological Mindedness as a Protective Factor Against Revictimization in Intimate Relationships. *Journal of Clinical Psychology*, *70*(9), 847–859. https://doi.org/10.1002/jclp.22061

Zorowitz, S., Niv, Y., & Bennett, D. (2021). *Inattentive responding can induce spurious associations between task behavior and symptom measures*. PsyArXiv. https://doi.org/10.31234/osf.io/rynhk

Zulueta, F. de. (2009). Post-traumatic stress disorder and attachment: Possible links with borderline personality disorder. *Advances in Psychiatric Treatment*, *15*(3), 172–180. https://doi.org/10.1192/apt.bp.106.003418

Appendix A: Attention Check Questions

To check for attentive responding, the following questions/statements, as adapted from Zorowitz et al. (2020), were utilized within the experiment:

1. I would feel bad if a loved one unexpectedly died (1 = “Not at all”; 5 = “Extremely”)
2. I would be able to lift a 1 lb (0.5 kg) weight (1 = “Never true”; 5 = “Very Often true”)
3. Over the last two weeks, how much time did you spend worrying about the 1977 Olympics? (1 = “None at all”; 5 = “A great deal”)
4. Have there been times of a couple days or more when you were able to stop breathing entirely (without the aid of medical equipment)? (1 = “None at all”; 5 = “A great deal”)

Attention checks 1 and 2 were imbedded within the prescreen measures while attention checks 3 and 4 were imbedded into the task itself.

Appendix B: Moral Preferences

Trial	Phase	κ	Δs	Δm	[s-, m-]	[s+, m+]	V_{harm}		P_{harm}		u^a	
							Good	Bad	Good	Bad	Good	Bad
1	Formation	0.5	17.125	17.125	[1,2.5]	[18,19.6]	-6.85	6.85	3.19E-298	1	0	1
2	Formation	0.814	2.625	11.517	[15,5.9]	[18,17.4]	1.617	7.274	1	1	1	1
3	Formation	0.399	14.28	9.486	[3,9.4]	[17,18.9]	-7.15	2.356	0	1	0	1
4	Formation	0.81	2.328	9.918	[7,4.9]	[9,14.8]	1.346	6.244	1	1	1	1
5	Formation	0.331	12.572	6.221	[5,9]	[18,15.2]	-6.934	0.583	7.08E-302	1	0	1
6	Formation	0.308	16.654	7.417	[1,12]	[18,19.4]	-9.433	0.196	0	0.999	0	1
7	Formation	0.375	1.941	1.166	[4,1.5]	[6,2.7]	-1.009	0.234	1.49E-44	1	0	1
8	Formation	0.601	10.773	16.217	[9,1]	[20,17.2]	-2.676	8.12	6.18E-117	1	0	1
9	Formation	0.898	1.853	16.34	[2,2.8]	[4,19.1]	3.605	10.882	1	1	1	1
10	Formation	0.271	10.296	3.829	[5,5.5]	[15,9.3]	-6.059	-0.408	7.57E-264	1.82E-18	0	0
11	Formation	0.688	6.731	14.864	[4,4.6]	[11,19.5]	-0.252	8.386	1.11E-11	1	0	1
12	Formation	0.625	10.299	17.15	[8,0.9]	[18,18]	-2.065	8.915	2.18E-90	1	0	1
13	Formation	0.186	9.137	2.083	[7,2.6]	[16,4.7]	-5.771	-1.283	2.34E-251	1.89E-56	0	0
14	Formation	0.332	5.573	2.773	[12,10.1]	[18,12.9]	-3.069	0.27	5.20E-134	1	0	1
15	Formation	0.737	5.057	14.189	[1,0.8]	[6,15]	0.717	8.415	1	1	1	1
16	Formation	0.713	4.304	10.694	[11,0.9]	[15,11.6]	0.196	6.195	0.999	1	1	1
17	Formation	0.469	4.157	3.669	[2,8.1]	[6,11.8]	-1.809	1.322	2.76E-79	1	0	1
18	Formation	0.19	13.24	3.108	[5,12.3]	[18,15.4]	-8.335	-1.796	0	9.59E-79	0	0
19	Formation	0.787	4.48	16.519	[15,2.3]	[19,18.8]	1.82	10.219	1	1	1	1
20	Formation	0.287	4.516	1.817	[15,2.2]	[20,4]	-2.616	-0.083	2.48E-114	2.58E-4	0	0
21	Formation	0.102	17.897	2.029	[2,8.4]	[20,10.4]	-11.919	-3.948	0	3.31E-172	0	0
22	Formation	0.531	5.012	5.678	[5,8.4]	[10,14.1]	-1.805	2.471	3.99E-79	1	0	1
23	Formation	0.668	9.699	19.489	[8,0.4]	[18,19.9]	-0.943	10.733	1.16E-41	1	0	1
24	Formation	0.263	6.137	2.187	[4,15.9]	[10,18.1]	-3.64	-0.31	8.25E-159	3.43E-14	0	0
25	Formation	0.312	3.454	1.564	[10,8.4]	[13,10]	-1.949	0.059	2.33E-85	0.997	0	1
26	Formation	0.731	4.44	12.085	[10,2.1]	[14,14.2]	0.518	7.128	1	1	1	1
27	Formation	0.213	16.306	4.422	[2,0.3]	[18,4.7]	-10.088	-1.796	0	9.54E-79	0	0
28	Formation	0.729	6.904	18.563	[4,1.3]	[11,19.9]	0.736	10.923	1	1	1	1

29	Formation	0.692	2.588	5.812	[4,2.3]	[7,8.1]	-0.068	3.292	1.08E-3	1	0	1
30	Formation	0.269	16.664	6.122	[2,1.1]	[19,7.2]	-9.828	-0.714	0	1.02E-31	0	0
31	Violation	0.501	17.438	17.499	[1,1.3]	[18,18.8]	7.018	-6.957	1	7.56E-303	1	0
32	Violation	0.707	3.503	8.433	[14,3.1]	[18,11.5]	4.852	0.078	1	0.999	1	1
33	Violation	0.589	9.913	14.193	[3,4.6]	[13,18.8]	6.961	-2.681	1	3.74E-117	1	0
34	Violation	0.873	2.776	18.998	[10,0.7]	[13,19.7]	12.466	3.756	1	1	1	1
35	Violation	0.133	17.719	2.727	[2,5.3]	[20,8]	-3.407	-11.586	1.06E-148	0	0	0
36	Violation	0.356	17.015	9.424	[1,0.6]	[18,10]	1.492	-9.083	1	0	1	0
37	Violation	0.127	9.18	1.341	[6,13.6]	[15,14.9]	-1.815	-6.024	1.47E-79	2.42E-262	0	0
38	Violation	0.132	17.457	2.662	[2,0.8]	[19,3.5]	-3.374	-11.422	2.98E-147	0	0	0
39	Violation	0.361	13.994	7.916	[6,6.7]	[20,14.6]	1.343	-7.421	1	0	1	0
40	Violation	0.868	2.427	15.919	[6,4]	[8,19.9]	10.415	3.077	1	1	1	1
41	Violation	0.867	1.189	7.724	[13,0.7]	[14,8.4]	5.05	1.485	1	1	1	1
42	Violation	0.55	13.155	16.086	[1,1.8]	[14,17.9]	7.314	-4.383	1	4.62E-191	1	0
43	Violation	0.462	12.804	10.981	[3,2.4]	[16,13.4]	3.846	-5.668	1	6.79E-247	1	0
44	Violation	0.397	9.353	6.15	[5,2.2]	[14,8.3]	1.499	-4.702	1	6.11E-205	1	0
45	Violation	0.207	11.679	3.056	[2,1.2]	[14,4.3]	-1.365	-7.259	5.33E-60	0	0	0
46	Violation	0.293	8.229	3.418	[5,10.5]	[13,13.9]	-0.076	-4.735	4.90E-4	2.27E-206	0	0
47	Violation	0.538	2.168	2.528	[9,10.6]	[11,13.1]	1.119	-0.759	1	1.05E-33	1	0
48	Violation	0.152	18.036	3.232	[1,12.8]	[19,16]	-3.149	-11.656	1.82E-137	0	0	0
49	Violation	0.411	11.203	7.824	[5,2.9]	[16,10.7]	2.116	-5.495	1	2.26E-239	1	0
50	Violation	0.065	10.229	0.707	[8,0.1]	[18,0.8]	-2.574	-6.948	1.62E-112	1.70E-302	0	0
51	Violation	0.848	1.764	9.846	[5,5]	[7,14.8]	6.363	1.719	1	1	1	1
52	Violation	0.603	6.4	9.733	[3,6.2]	[9,15.9]	4.893	-1.56	1	1.80E-68	1	0
53	Violation	0.45	2.16	1.766	[4,7.2]	[6,9]	0.588	-0.982	1	2.24E-43	1	0
54	Violation	0.793	3.967	15.164	[13,2]	[17,17.2]	9.424	1.772	1	1	1	1
55	Violation	0.644	1.265	2.284	[7,5.4]	[8,7.7]	1.219	-0.2	1	2.00E-09	1	0
56	Violation	0.935	1.074	15.546	[12,3.8]	[13,19.3]	10.56	3.912	1	1	1	1
57	Violation	0.295	12.889	5.383	[4,8.1]	[17,13.5]	-0.098	-7.407	5.36E-05	0	0	0
58	Violation	0.639	1.78	3.146	[4,11.2]	[6,14.3]	1.668	-0.302	1	7.73E-14	1	0
59	Violation	0.499	6.17	6.149	[2,8.7]	[8,14.8]	2.453	-2.475	1	3.38E-108	1	0
60	Violation	0.705	2.626	6.286	[5,1]	[8,7.3]	3.613	0.048	1	0.992	1	1
61	Reparation	0.65	1.203	2.234	[5,8.2]	[6,10.4]	-0.172	1.203	3.32E-08	1	0	1

62	Reparation	0.722	6.157	15.998	[6,2.8]	[12,18.8]	0.49	9.352	1	1	1	1
63	Reparation	0.435	3.859	2.968	[5,2.6]	[9,5.6]	-1.811	0.92	2.20E-79	1	0	1
64	Reparation	0.704	2.379	5.661	[9,11.6]	[11,17.3]	0.033	3.249	0.964	1	1	1
65	Reparation	0.825	2.74	12.909	[17,2.7]	[20,15.6]	1.955	8.214	1	1	1	1
66	Reparation	0.242	5.102	1.625	[14,0.8]	[19,2.4]	-3.084	-0.393	1.18E-134	8.30E-18	0	0
67	Reparation	0.278	6.368	2.451	[6,15.2]	[12,17.7]	-3.722	-0.195	2.16E-162	3.42E-09	0	0
68	Reparation	0.851	3.3	18.809	[2,0.8]	[5,19.6]	3.332	12.176	1	1	1	1
69	Reparation	0.296	1.798	0.757	[17,17.2]	[19,18]	-1.031	-0.009	1.65E-45	0.287	0	0
70	Reparation	0.753	5.45	16.649	[6,2.7]	[11,19.3]	1.18	10.019	1	1	1	1
71	Reparation	0.63	3.273	5.568	[5,10.2]	[8,15.8]	-0.62	2.916	1.13E-27	1	0	1
72	Reparation	0.653	9.229	17.385	[10,1.1]	[19,18.5]	-1.245	9.401	8.44E-55	1	0	1
73	Reparation	0.387	19.296	12.189	[1,5.8]	[20,18]	-9.85	2.743	0	1	0	1
74	Reparation	0.382	18.715	11.552	[1,0.5]	[20,12.1]	-9.635	2.472	0	1	0	1
75	Reparation	0.347	12.051	6.398	[2,13.4]	[14,19.8]	-6.517	0.863	9.81E-284	1	0	1
76	Reparation	0.149	6.26	1.098	[5,5.5]	[11,6.6]	-4.052	-1.109	1.04E-176	6.85E-49	0	0
77	Reparation	0.794	2.667	10.287	[7,7.2]	[10,17.5]	1.219	6.401	1	1	1	1
78	Reparation	0.618	10.603	17.179	[4,2.5]	[15,19.7]	-2.269	8.844	2.92E-99	1	0	1
79	Reparation	0.296	12.845	5.398	[1,7.4]	[14,12.8]	-7.372	-0.075	0	5.61e-4	0	0
80	Reparation	0.758	5.556	17.446	[1,1.6]	[7,19]	1.345	10.545	1	1	1	1
81	Reparation	0.37	3.18	1.869	[15,2.6]	[18,4.5]	-1.665	0.354	4.81E-73	1	0	1
82	Reparation	0.175	1.814	0.385	[15,0.9]	[17,1.3]	-1.154	-0.275	7.45E-51	1.17E-12	0	0
83	Reparation	0.206	18.331	4.753	[1,5.5]	[19,10.3]	-11.406	-2.172	0	4.65E-95	0	0
84	Reparation	0.565	6.337	8.24	[10,6.2]	[16,14.4]	-1.964	3.867	5.20E-86	1	0	1
85	Reparation	0.613	7.431	11.763	[1,0.6]	[8,12.4]	-1.672	6.005	2.33E-73	1	0	1
86	Reparation	0.247	14.363	4.701	[5,4.9]	[19,9.6]	-8.644	-1.018	0	6.19E-45	0	0
87	Reparation	0.704	1.456	3.456	[12,14]	[13,17.5]	0.018	1.982	0.852	1	1	1
88	Reparation	0.35	14.763	7.953	[3,0.2]	[18,8.2]	-7.948	1.138	0	1	0	1
89	Reparation	0.483	6.069	5.672	[7,1]	[13,6.7]	-2.547	2.149	2.53E-111	1	0	1
90	Reparation	0.5	16.753	3	[1.6,20]	[18.4,0.5]	0	1	0	1	0	1

[a] If u = 1, the agent choose the harmful option (C_{harm}). If u = 0, the agent choose the least harmful option (C_{help}).

Appendix C: Supplementary Analysis with Childhood Trauma Group

This supplementary analysis duplicates the results from the main analysis with the caveat that the Childhood Trauma group has been included in the analysis (n = 13). Trends observed in the main analysis were replicated within this analysis, suggesting robustness. However, the addition of the Childhood Trauma group also saw to the detection of an additional significant Group x Phase x Agent interaction for impression ratings as well as differences in the post-hoc analysis for the Group x Agent interaction for certainty ratings. Data analysis used in the supplementary analysis were largely the same, with some exceptions. Instead of a Mann Whitney U-test for the prescreen measures, a Kruskal-Wallis Rank Sum Test was run instead, as it is the equivalent of to a one-way ANOVA. Post-hoc analysis was conducted for the Kruskal Wallis in the form of pairwise comparisons via the Dunn's test. Similarly, for demographics, Fisher's exact test was conducted to assess for differences between the three groups, with pairwise comparisons also being conducted using Fisher's exact test to compare two groups at a time.

Table 1

ANOVA Results for Impression

Measure	*df*	*F*	η_p^2	*p* value	
Group	2, 82	1.72	.04	.186	
Phase	1.6, 131.3	2.57	.03	.092	.
Group : Phase	3.2, 131.3	0.27	.006	.862	
Agent	1, 82	68.7	.456	<.001	***
Group : Agent	2, 82	2.43	.056	.095	.
Phase : Agent	1, 94	20.4	.199	<.001	***
Group : Phase : Agent	3.88, 159.3	3.64	.082	.008	**

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes (.01 < η_p^2 < .06) for Group, Phase, and Group x Agent. Medium effects (.06 < η_p^2 < .14) were present for Group x Phase x Agent. Large effects (η_p^2 > .14) were present for Agent and Phase x Agent.

Figure 1

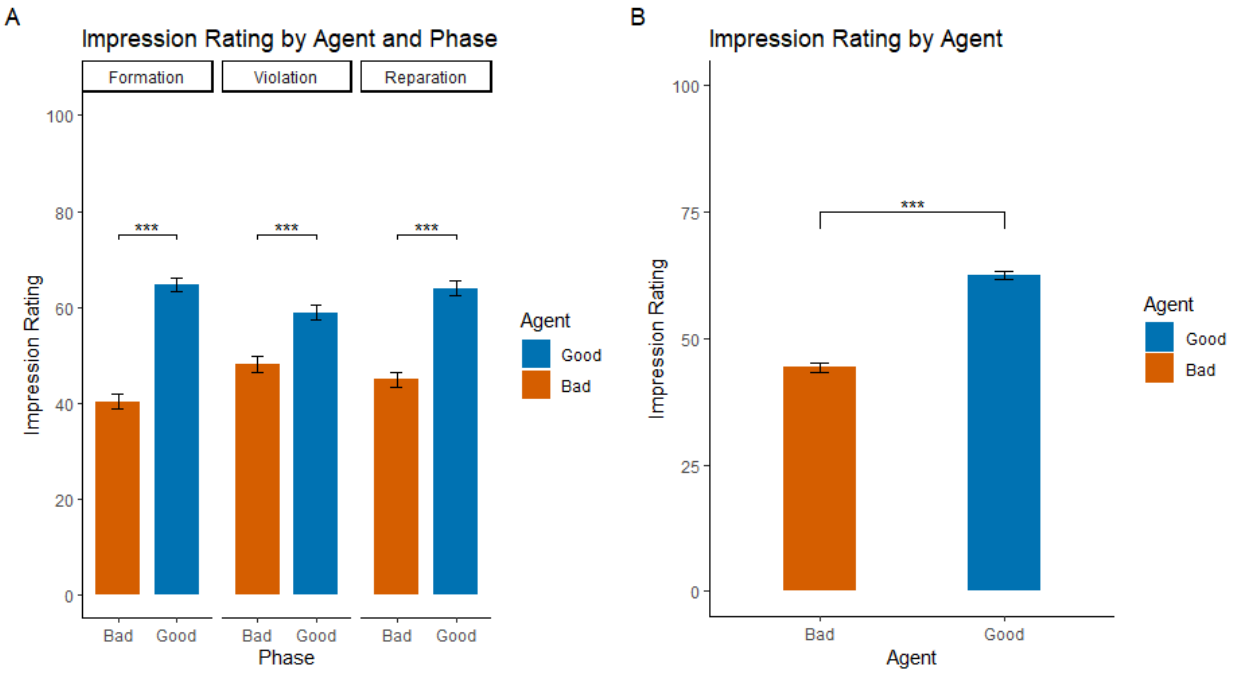
Impression Ratings by Agent and Group throughout Task, including Significance

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Majority of the trends observed in the main analysis were still present even after the inclusion of the Childhood Trauma group, with the exception that there was no longer a marginally significant effect for Group. The inclusion of the Childhood Trauma group, however, led to the detection of a significant three-way interaction (Group x Phase x Agent) that was not present within the main analysis, $F(3.88, 159.3) = 3.64$, $p = .008$, $\eta_p^2 = .082$. It also detected a marginally significant interaction for Group x Agent, $F(2, 82) = 20.4$, $p = .095$, $\eta_p^2 = .095$.

Figure 2

Significant Main Effects and Interactions for Impression Rating



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 2

ART ANOVA Results for Certainty

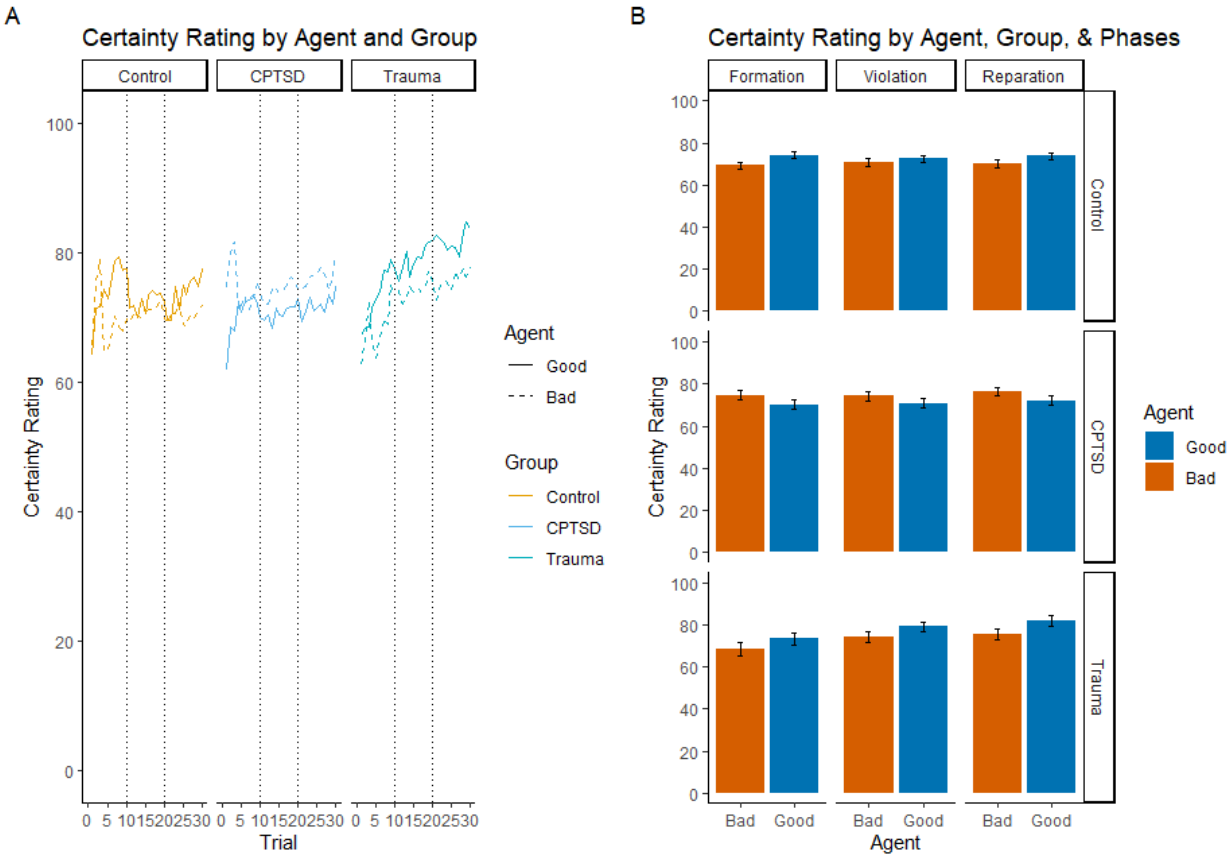
Measure	*df*	*F*	η_p^2	*p* value	
Group	2, 82	0.257	.006	.774	
Phase	2, 410	1.33	.006	.266	
Group : Phase	4, 410	1.82	.017	.124	
Agent	1, 410	0.479	.001	.489	
Group : Agent	2, 410	10.8	.05	<.001	***
Phase : Agent	2, 410	0.129	.001	.879	
Group : Phase : Agent	4, 410	0.404	.004	.806	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) indicated a small effect size (.01 < η_p^2 < .06) for Group x Phase and Group x Agent.

Results for certainty ratings in the ANOVA were consistent with the results observed in the main analysis, even with the inclusion of the Childhood Trauma group. However, the results of the post-hoc analysis differed as it did not detect any significant differences between groups.

Figure 3

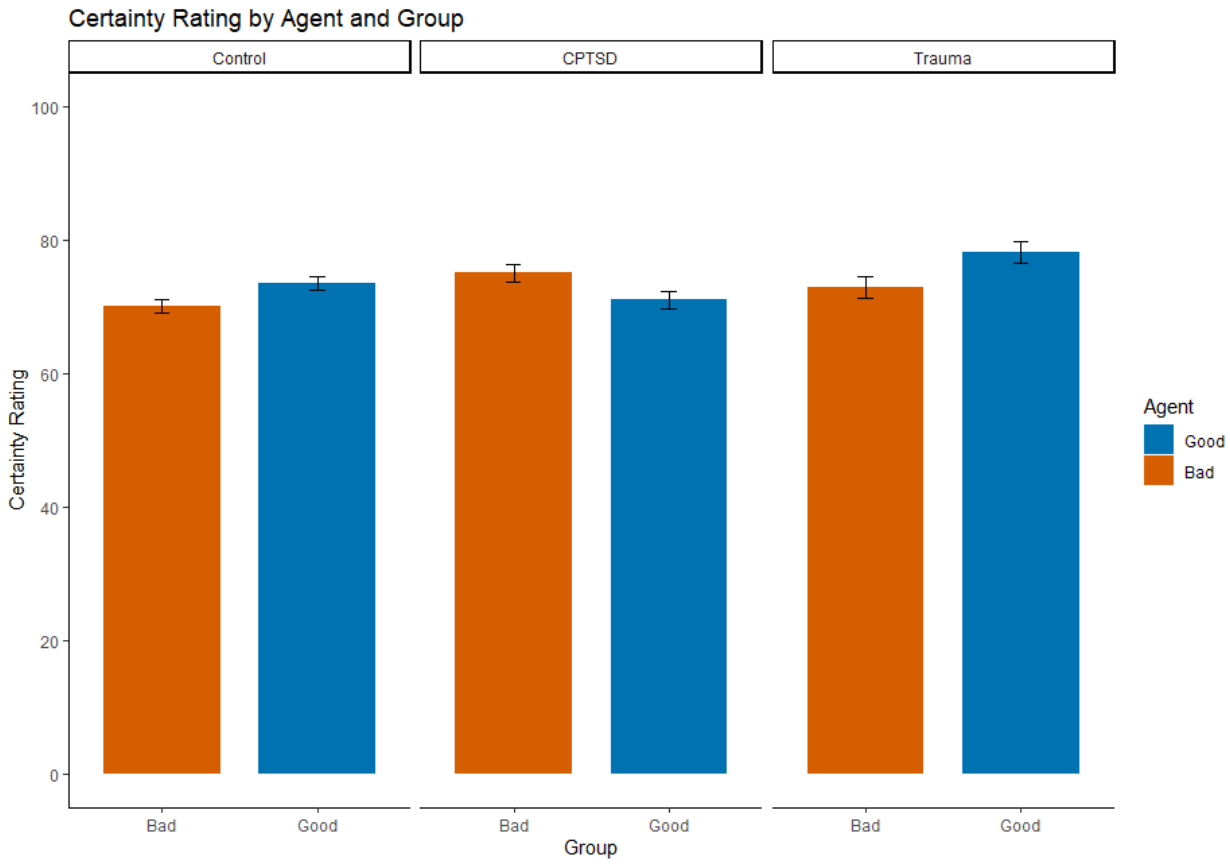
Certainty Ratings by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Figure 4

Significant Main Interaction for Certainty Ratings



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 3

ART ANOVA Results for Prediction Accuracy

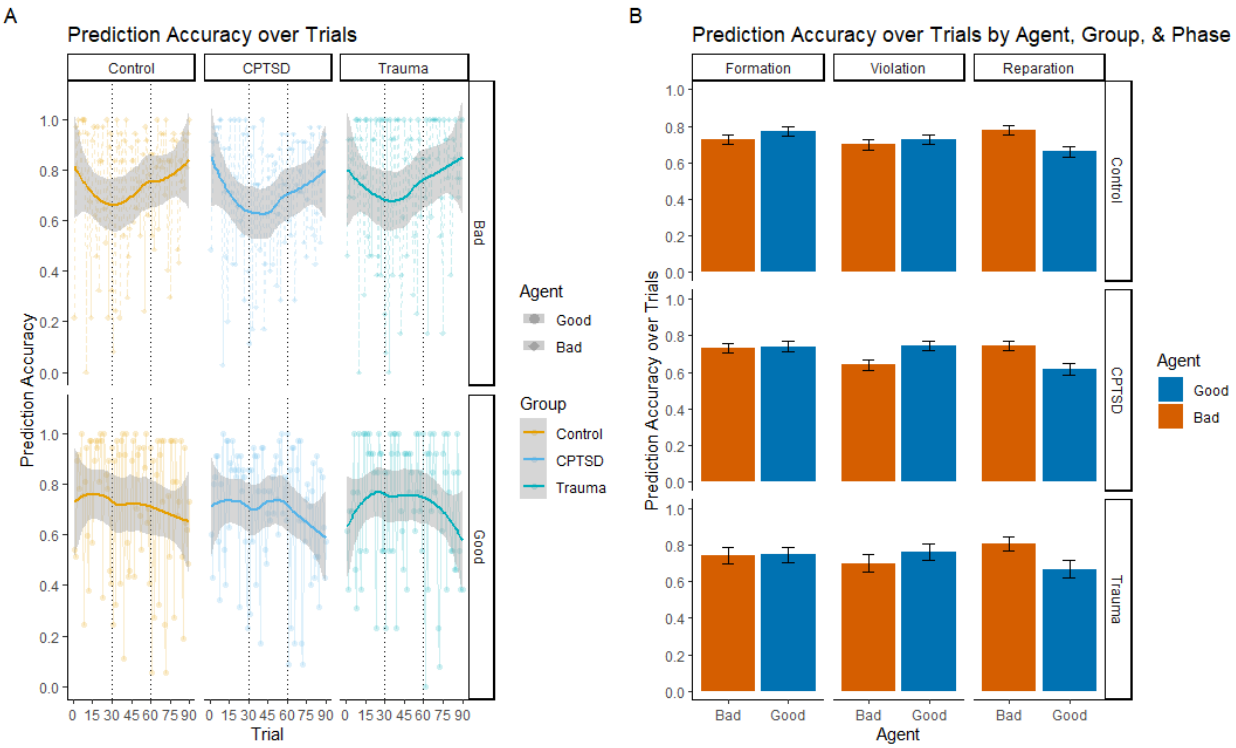
Measure	*df*	*F*	η_p^2	*p* value	
Group	2, 82	1.96	.046	.147	
Phase	2, 410	6.95	.033	.001	**
Group : Phase	4, 410	0.759	.007	.552	
Agent	1, 410	1.68	.004	.196	
Group : Agent	2, 410	0.488	.002	.614	
Phase : Agent	2, 410	39.55	.162	<.001	***
Group : Phase : Agent	4, 410	1.25	.012	.287	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes (.01 < η_p^2 < .06) for Group, Phase, and Group x Phase x Agent. Large effects (η_p^2 <.14) were present for Phase x Agent.

Results for prediction accuracy were consistent with the results observed in the main analysis, even with the inclusion of the Childhood Trauma group. No notable differences were detected.

Figure 5

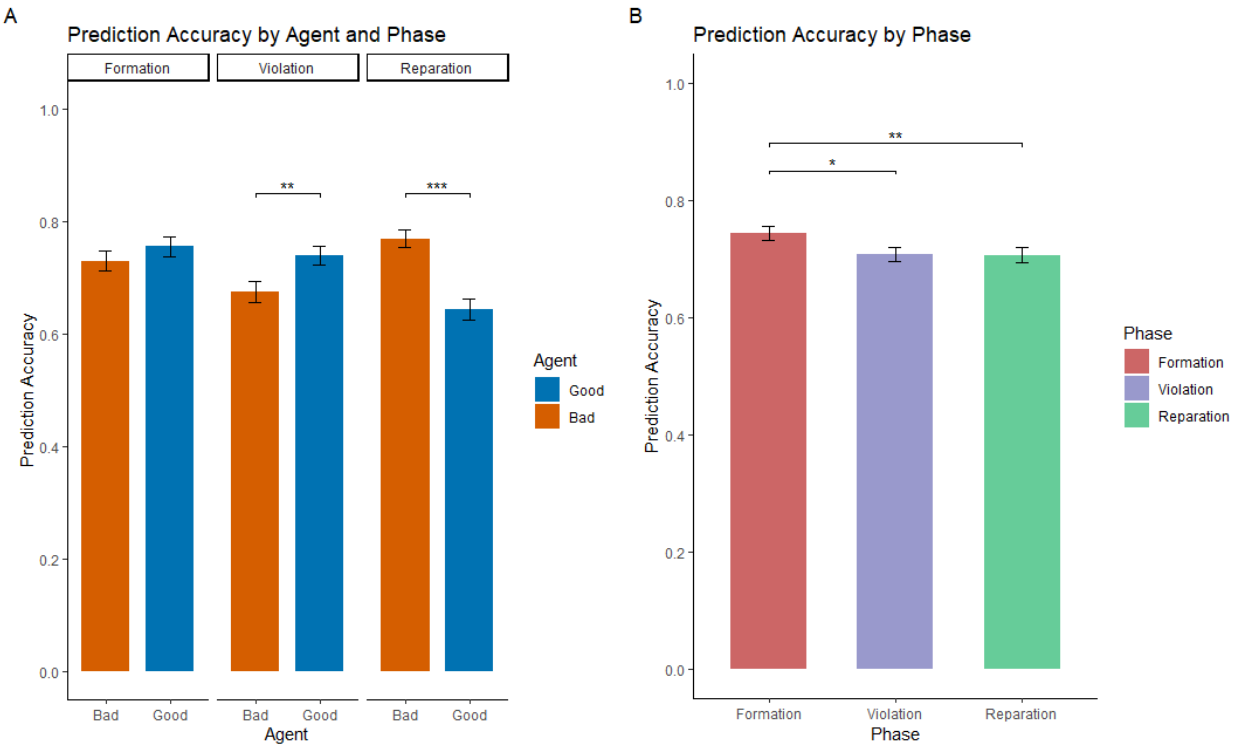
Prediction Accuracy by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Average Prediction Accuracy per Trial was mapped in the background and a local polynomial regression (solid line) was fitted over it. Error bands represent 95% CI. (B) Error bars represent 95% CI.

Figure 6

Significant Main Effects and Interactions for Prediction Accuracy



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 4

Pre-Screen Trauma Surveys

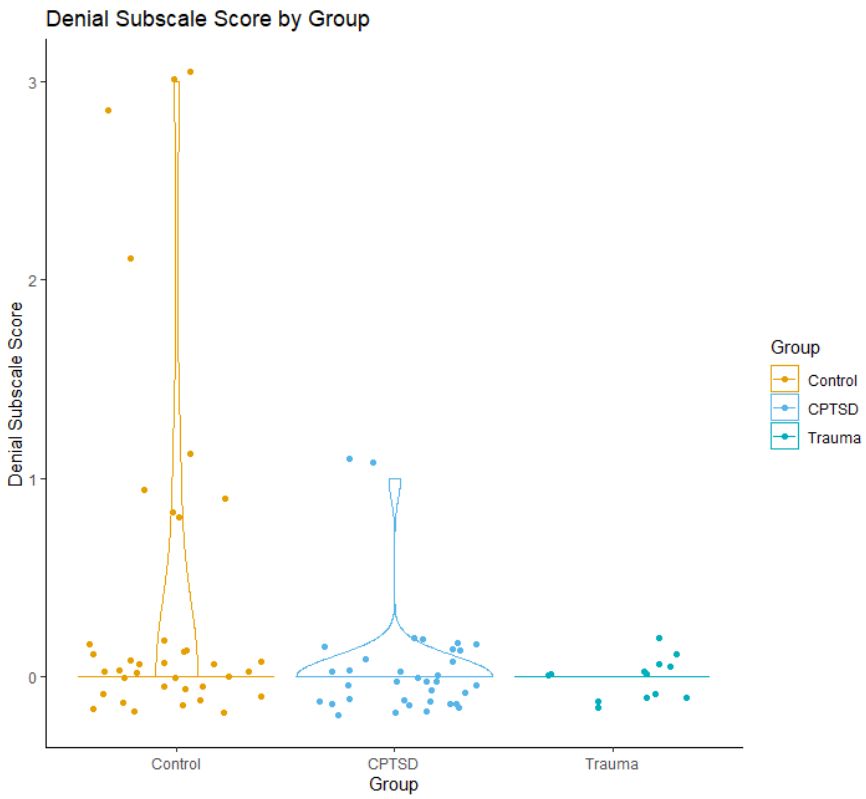
	Childhood Trauma (n = 13, 15.29%)	CPTSD (n = 35, 41.18%)	Control (n = 37, 43.52%)	Overall (n = 85)	χ^2 (2)	*p* value
ACEs	3	5	0	3	53.2	<.001
CTQ	56	68	40	48	54.7	<.001
Emotional Abuse	13	18	6	8	52.4	<.001
Emotional Neglect	14	18	8	12	45.8	<.001
Physical Abuse	7	8	5	6	29.6	<.001
Physical Neglect	10	11	5	7	51.0	<.001
Sexual Abuse	9	8	5	5	31.7	<.001
Denial[a]	0	0	0	0	7.99	.018

Note. Values reflect medians due to non-parametric distribution. *p* values were derived from Kruskal-Wallis Rank Sum Tests that were conducted to compare CPTSD, Control, and Childhood Trauma groups. See Figure 8 for Post-hoc analysis. Pairwise comparisons were performed using Dunn's test.

[a] Medians for subscale were identical but distribution differed. See Figure 7 for the direction of significance.

Figure 7

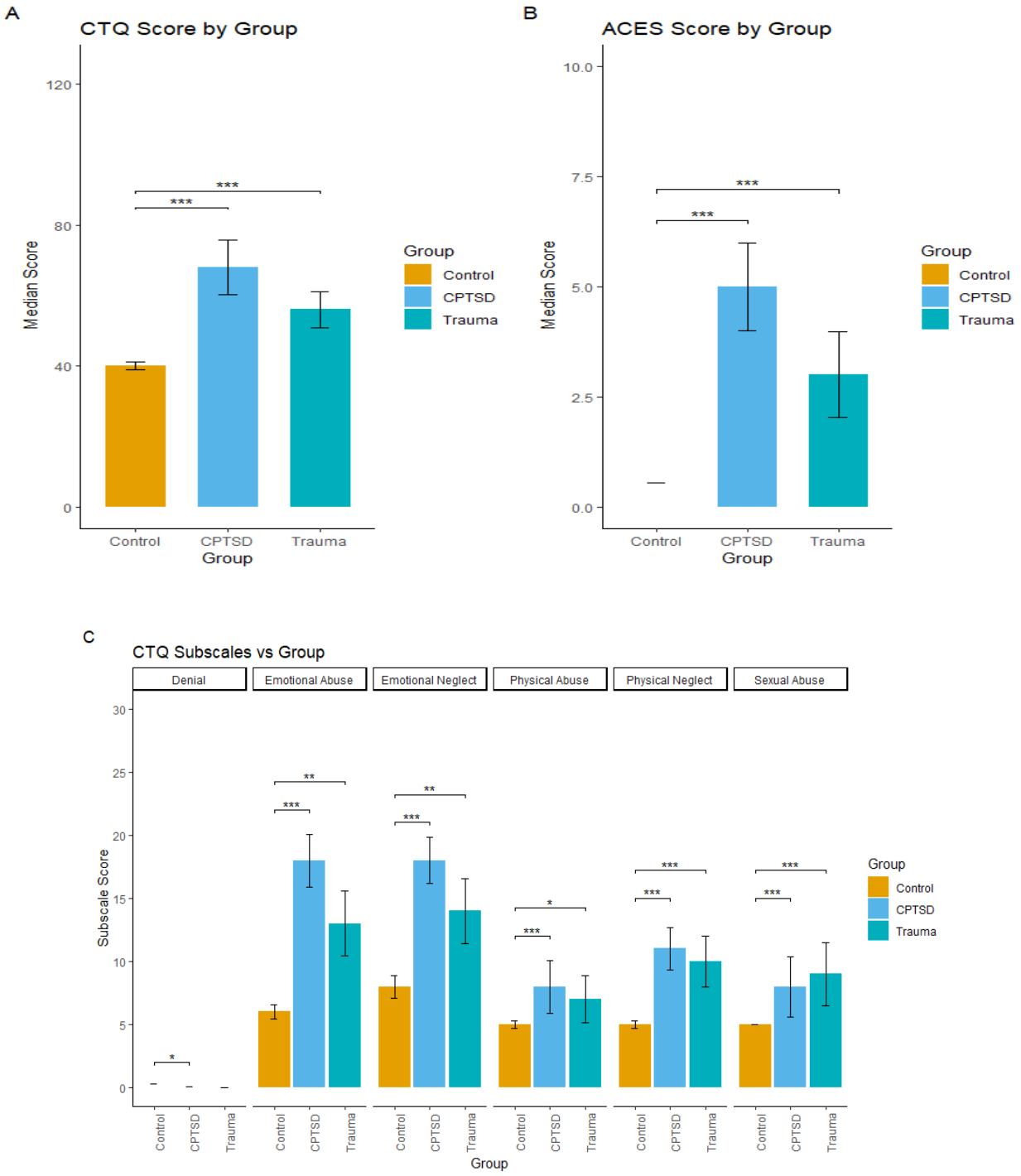
Distribution of Denial Subscale



Note. The CTQ Denial Subscale measures underreporting of trauma, with increasing minimization as the score increases, with a max score of 3 and a minimum score of 0. Data points were jittered to prevent overlap—any points below 0 on the graph indicate a denial subscale score of 0.

Figure 8

Post-hoc Analysis for Pre-Screen Surveys



Note. *p<.05, **p<.01, ***p<.001.

Table 5

Participant Demographics

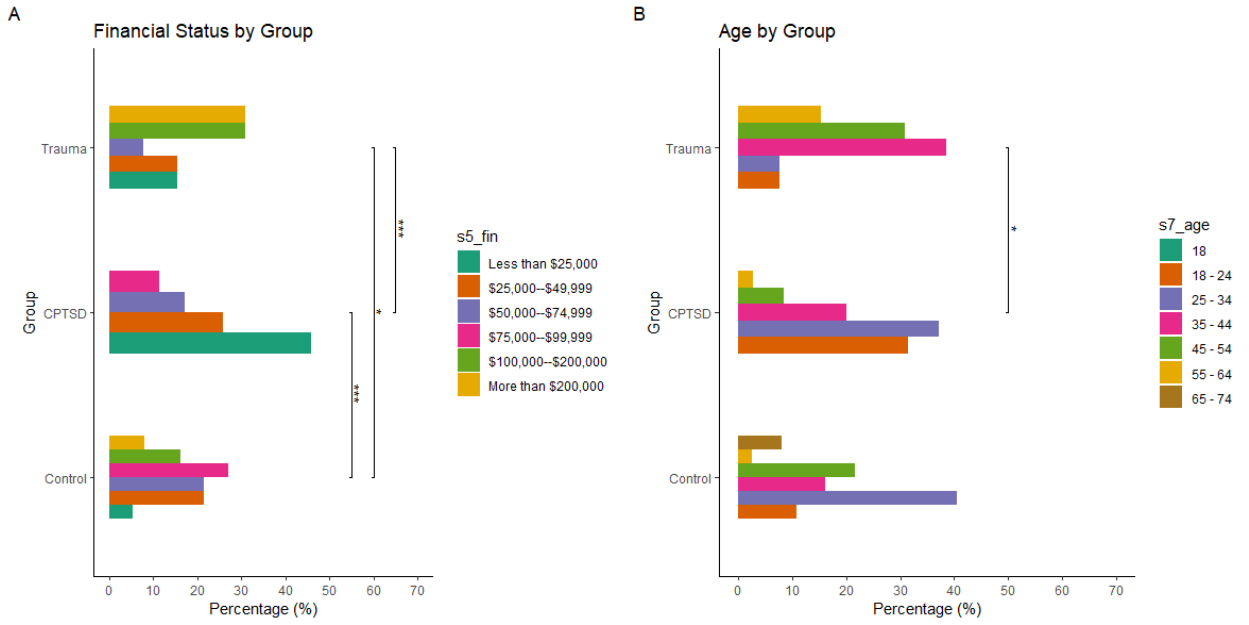
	Childhood Trauma (n = 13, 15.29%)		CPTSD (n = 35, 41.18%)		Control (n = 37, 43.52%)		Overall (n = 61)		*p* value
	n	%	n	%	n	%	n	%	
Race									.377
White	9	69.2	26	74.3	29	78.4	64	75.4	
Asian	2	15.4	1	2.9	5	13.5	8	9.4	
American Indian or Alaskan Native	-	-	2	5.7	-	-	2	2.4	
Multiple Selected	2	15.4	5	14.3	3	8.1	10	11.8	
Other	-	-	1	2.9	-	-	1	1.2	
Education									.451
High school graduate or equivalent	1	7.7	4	11.4	2	5.4	7	8.2	
Some college but not degree	3	23.1	13	37.1	9	24.3	25	29.4	
Associate degree	1	7.7	3	8.6	1	2.7	5	5.9	
Bachelor’s degree	7	53.8	11	31.4	16	43.2	34	40	
Master’s degree	-	-	4	11.4	7	18.9	11	12.9	
Doctoral degree	-	-	-	-	1	2.7	1	1.2	
Professional degree	1	7.7	-	-	1	2.7	2	2.4	
Gender									.450
Female	7	53.8	23	65.7	23	62.2	53	62.4	
Male	6	46.2	9	25.7	14	37.8	29	34.1	
Gender Nonbinary	-	-	2	5.7	-	-	2	2.4	
Other	-	-	1	2.9	-	-	1	1.2	
Relationship Status									.121
Never Married	5	38.5	27	77.1	22	59.5	54	63.5	
Married	6	46.2	7	20	12	32.4	25	29.4	
Divorced	2	15.4	1	2.9	2	5.4	5	5.9	
Widowed	-	-	-	-	1	2.7	1	1.2	
Financial Status									<.001[a]
Less than $25,000	2	15.4	16	45.7	2	5.4	20	23.5	
$25,000 - $49,999	2	15.4	9	25.7	8	21.6	19	22.4	
$50,000 - $74,999	1	7.7	6	17.1	8	21.6	15	17.6	
$75,000 - $99,999	-	-	4	11.4	10	27	14	16.5	
$100,000 - $200,000	4	30.8	-	-	6	16.2	10	11.8	
More than $200,000	4	30.8	-	-	3	8.1	7	8.2	
Age									.015[a]

18 – 24	1	7.7	11	31.4	4	10.8	16	18.8
25 – 34	1	7.7	13	37.1	15	40.5	29	34.1
35 – 44	5	38.5	7	20	6	16.2	18	21.2
45 – 54	4	30.8	3	8.6	8	21.6	15	17.6
55 – 64	2	15.4	1	2.9	1	2.7	4	4.7
65 – 74	-	-	-	-	3	8.1	3	3.5

[a] See Figure 9 for post-hoc analysis, conducted using Fisher's Exact Test.

Figure 9

Pairwise Fisher's Exact Test for Age and Financial Status



Note. *p<.05, **p<.01, ***p<.001.

For the prescreen measures, it was generally found that there were significant differences in results between the Control group and the CPTSD group, as well as significant differences between the Control group and the Childhood Trauma group. The only exception to this was with the Denial subscale, where the only significance detect was between the Control group and the CPTSD group. Differences between the CPTSD group and Childhood Trauma were largely insignificant. In terms of demographics, the groups primarily differed by financial status and age.

All three groups were significant different from each other by financial status, with the CPTSD group typically occupying a lower income bracket. The Control group appeared to be roughly equal split in terms of financial status while those with Childhood Trauma were largely in higher income brackets. In terms of age, there were significant differences between the CPTSD group and the Childhood Trauma group, with the former occupying younger age groups.

Appendix D: Supplementary Analysis with Maximum Denial Subscale Score Removal

This supplementary analysis duplicates the results from the main analysis with the caveat that participants who reported a denial subscale score of 3 have been removed from the analysis (n = 3). Analysis is largely in line with the results reported in the main analysis, suggesting that the trends reported in the main analysis are fairly robust. Additional significant were found however, specifically a significant Group x Agent interaction for impression rating and a Group x Phase x Agent interaction for prediction accuracy. Data analysis used in the supplementary analysis was virtually identical to that of the main analysis.

Table 1

ART ANOVA Results for Impression

Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 67	3.91	.055	.052	.
Phase	2, 335	1.20	.007	.302	
Group : Phase	2, 335	0.099	.001	.906	
Agent	1, 335	285.2	.460	<.001	***
Group : Agent	1, 335	7.69	.022	.005	**
Phase : Agent	2, 335	16.5	.090	<.001	***
Group : Phase : Agent	2, 335	0.385	.003	.681	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes ($.01 < \eta_p^2 < .06$) for Group and Group x Agent. Medium effects ($.06 < \eta_p^2 < .14$) were present for Phase x Agent. Large effects ($\eta_p^2 > .14$) were present for Agent.

Trends observed in the main analyses were largely preserved for impression ratings apart from two points. The marginal significant effect for Phase was no longer preserved after the omission and an additional significant Group x Agent interaction emerged, $F(1, 335) = 7.69$, $p = .005$, $\eta_p^2 = .022$, as well (see Figure 2C for post-hoc analysis). Notably, as the omission led to a violation of ANOVA assumptions, an ART ANOVA was conducted instead of a parametric Mixed ANOVA.

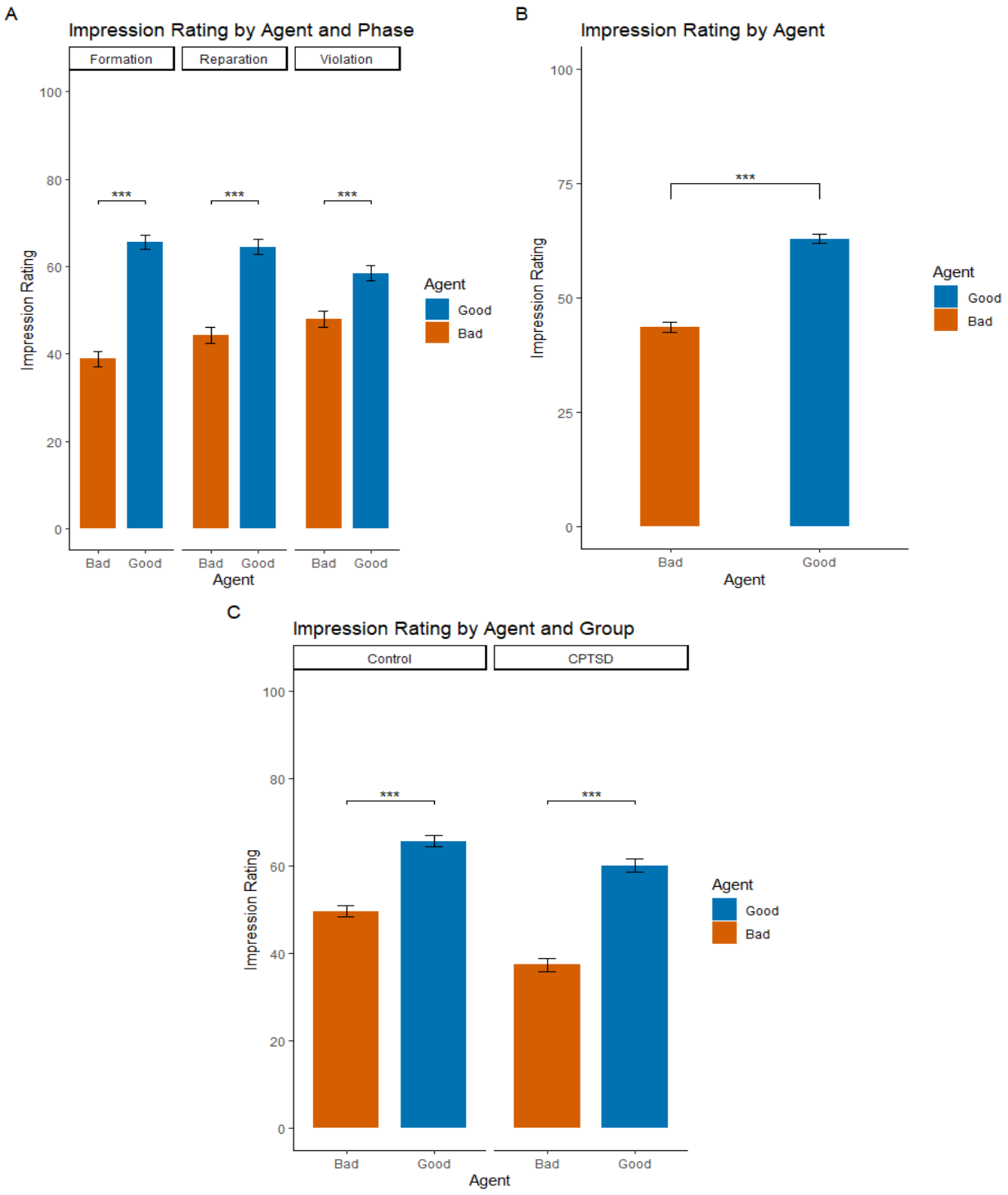
Figure 1

Impression Ratings by Agent and Group throughout Task

Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Figure 2

Significant Main Effects and Interactions for Impression Rating



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 2

ART ANOVA Results for Certainty

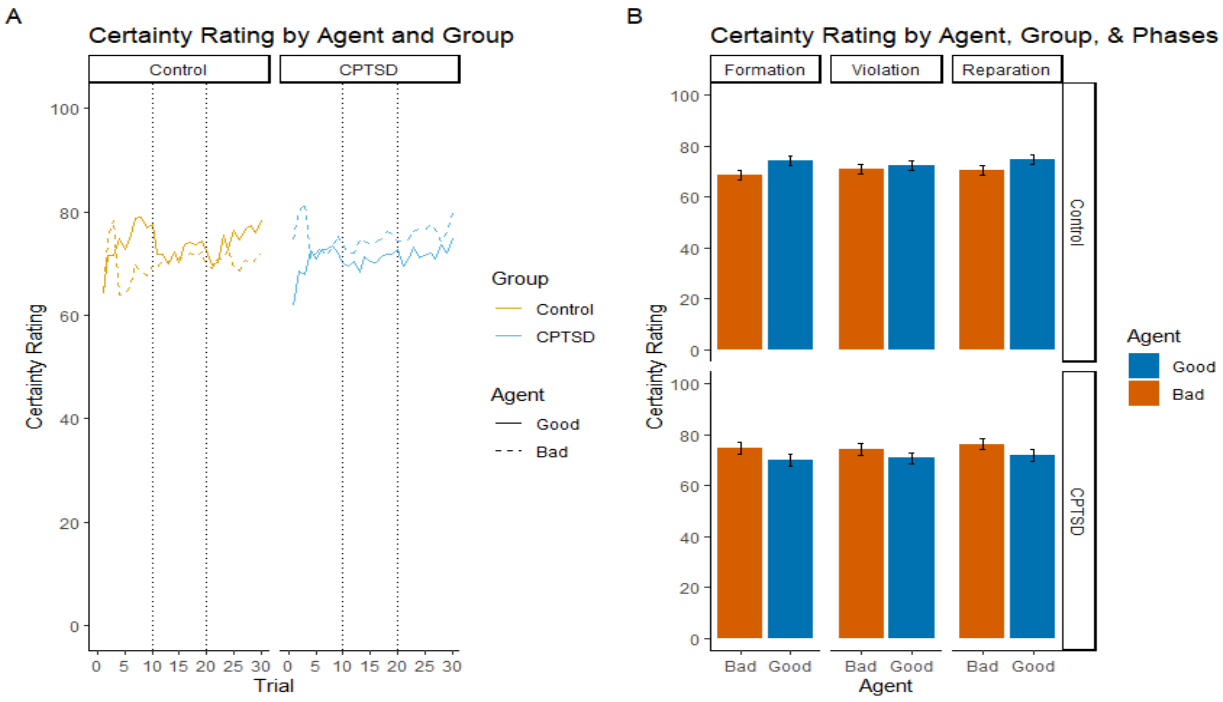
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 67	0.152	.002	.698	
Phase	2, 335	0.793	.004	.453	
Group : Phase	2, 335	0.066	<.001	.936	
Agent	1, 335	1.10	<.001	.748	
Group : Agent	1, 335	16.0	.046	<.001	***
Phase : Agent	2, 335	0.288	.002	.750	
Group : Phase : Agent	2, 335	0.968	.006	.381	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) indicated a small effect size ($.01 < \eta_p^2 < .06$) for Group x Agent.

Trends observed in the main analysis were replicated for certainty ratings, with no notable differences between the supplementary and main analysis.

Figure 3

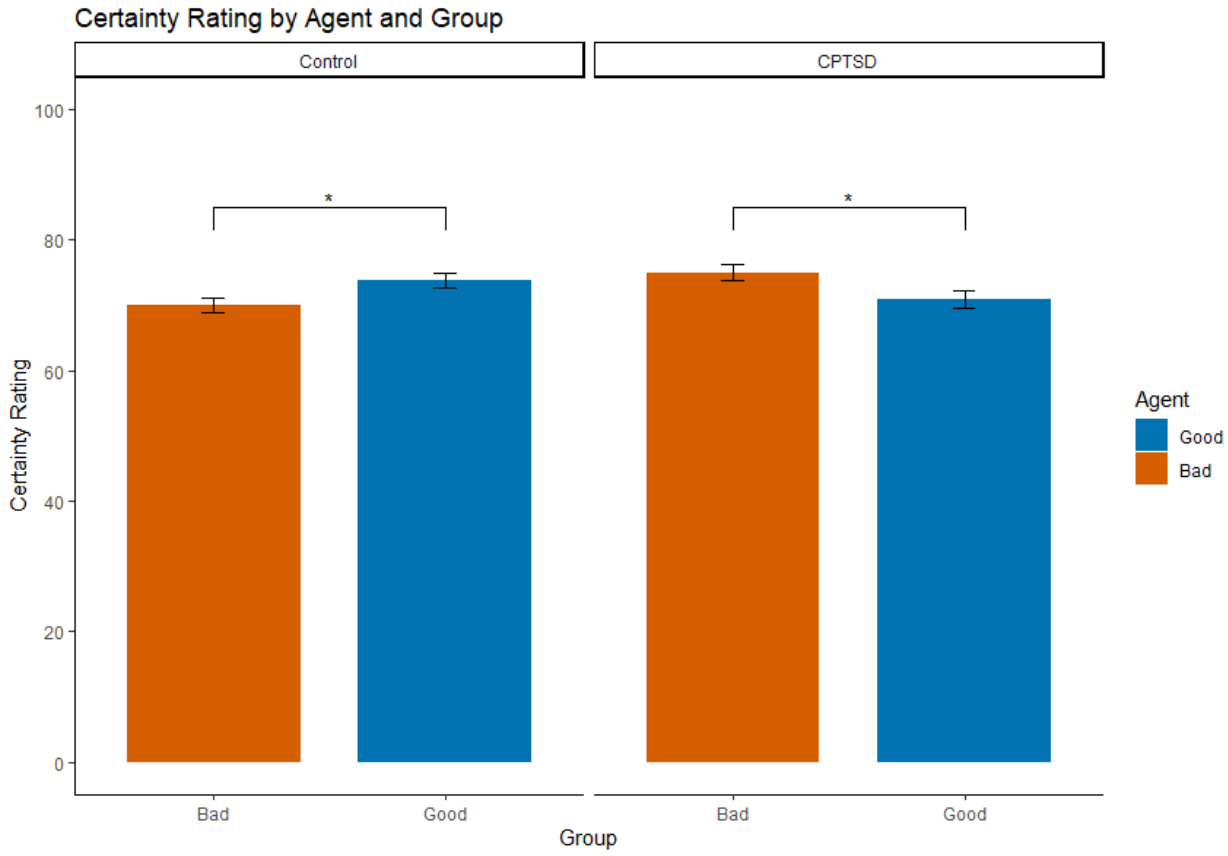
Certainty Ratings by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Figure 4

Significant Main Interaction for Certainty Ratings



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 3

ART ANOVA Results for Prediction Accuracy

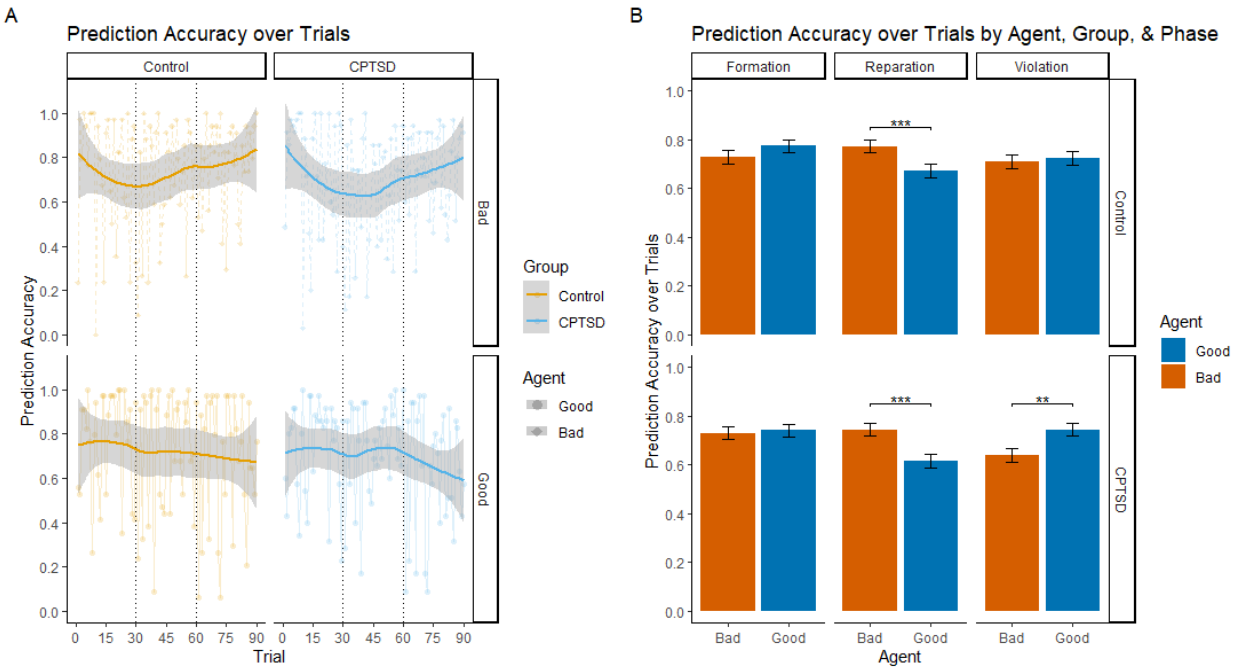
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 67	2.61	.038	.111	
Phase	2, 335	8.84	.05	<.001	***
Group : Phase	2, 335	0.458	.003	.633	
Agent	1, 335	1.49	.004	.224	
Group : Agent	1, 335	0.243	<.001	.622	
Phase : Agent	2, 335	35.9	.176	<.001	***
Group : Phase : Agent	2, 335	3.57	.021	.029	*

Note. · p<01, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes (.01 < η_p^2 < .06) for Group, Phase, and Group x Phase x Agent. Large effects (η_p^2 <.14) were present for Phase x Agent.

Trends observed in the main analysis for prediction accuracy were replicated in the supplementary analysis. The supplementary analysis additionally observed a significant interaction between Group x Phase x Agent, $F(2, 335) = 3.57$, $p = .029$, $\eta_p^2 = .021$.

Figure 5

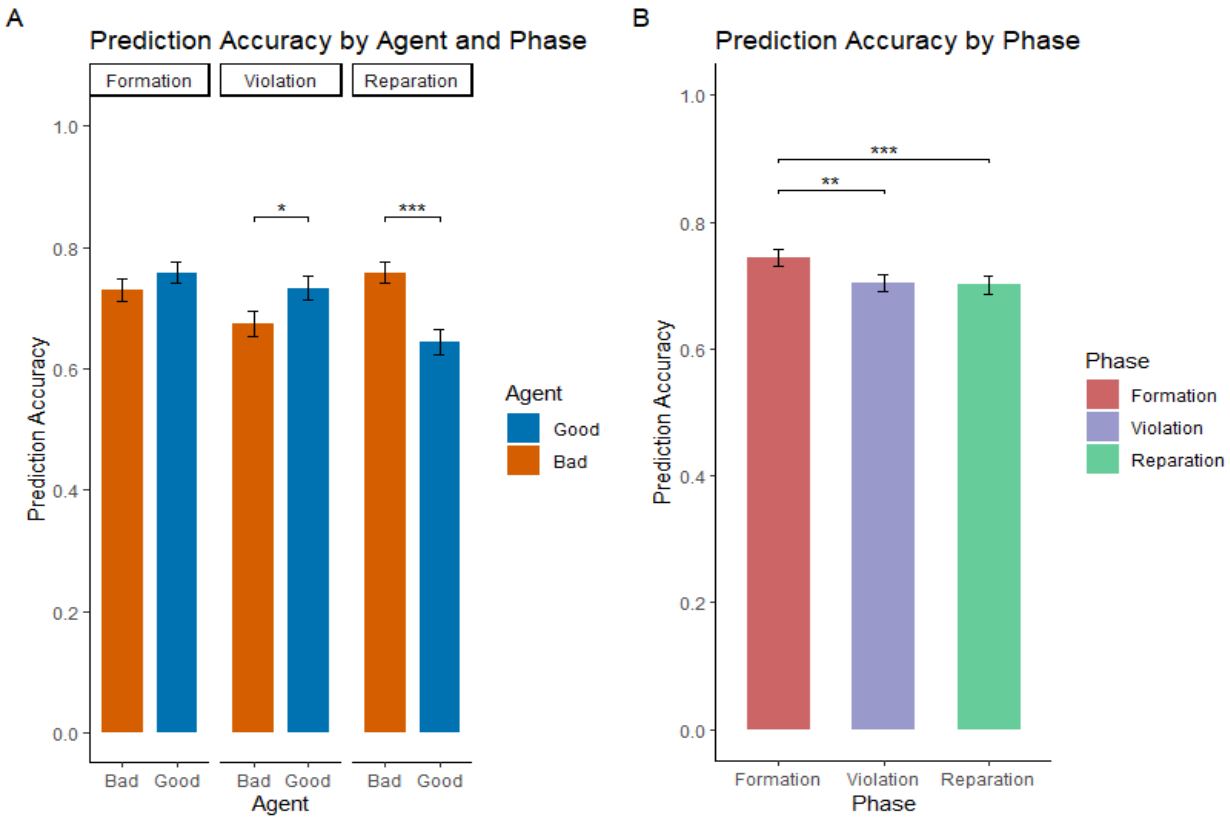
Prediction Accuracy by Group and Agent throughout Task, including Significance



Note. *p<.05, **p<.01, ***p<.001. (A) Dotted lines indicate the beginning and endings of various task phases. Average Prediction Accuracy per Trial was mapped in the background and a local polynomial regression (solid line) was fitted over it. Error bands represent 95% CI. (B) Error bars represent 95% CI.

Figure 6

Significant Main Effects and Interactions for Prediction Accuracy



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 4

Pre-Screen Trauma Surveys

	CPTSD (n = 35, 50.72%)	Control (n = 34 , 49.28%)	Overall (n = 69)	*U*	*p* value	r
ACEs	5	0	2	-6.60	<.001	.794
CTQ	68	40	46	-6.55	<.001	.789
Emotional Abuse	18	6	9	-6.63	<.001	.799
Emotional Neglect	18	8	12	-6.10	<.001	.734
Physical Abuse	8	5	6	-5.12	<.001	.616
Physical Neglect	11	5	7	-6.38	<.001	.768
Sexual Abuse	8	5	5	-5.26	<.001	.633
Denial	0	0	0	1.56	.126	.187

Note. Values reflect medians due to non-parametric distribution. *p* values were derived from Mann Whitney U Tests that were conducted to compare CPTSD and Control groups. Wilcoxon effect sizes (r) in these cases exhibit large effects ($r \geq .5$) except for the Denial subscale, which exhibited a small effect ($.1 < r < .3$)

Table 5

Participant Demographics

	CPTSD (n = 35, 50.72%)		Control (n = 34 , 49.28%)		Overall (n = 72)		*p* value
	n	%	n	%	n	%	
Race							.116
White	26	74.3	27	79.4	53	76.8	
Asian	1	2.9	5	14.7	6	8.7	
American Indian or Alaskan Native	2	5.7	-	-	2	2.9	
Multiple Selected	5	14.3	2	5.9	7	10.1	
Other	1	2.9	-	-	1	1.4	
Education							.325
High school graduate or equivalent	4	11.4	2	5.9	6	8.7	
Some college but not degree	13	37.1	7	20.6	20	29	
Associate degree	3	8.6	1	2.9	4	5.8	
Bachelor’s degree	11	31.4	15	44.1	26	37.7	
Master’s degree	4	11.4	7	20.6	11	15.9	
Doctoral degree	-	-	1	2.9	1	1.4	
Professional degree	-	-	1	2.9	1	1.4	
Gender							.295
Female	23	65.7	21	61.8	44	63.8	
Male	9	25.7	13	38.2	22	31.9	

Gender Nonbinary	2	5.7	-	-	2	2.9	
Other	1	2.9	-	-	1	1.4	
Relationship Status							.384
Never Married	27	77.1	20	58.8	47	68.1	
Married	7	20.0	11	32.4	18	26.1	
Divorced	1	2.9	2	5.9	3	4.3	
Widowed	-	-	1	2.9	1	1.4	
Financial Status							<.001
Less than $25,000	16	45.7	1	2.9	17	24.6	
$25,000 - $49,999	9	25.7	8	23.5	17	24.6	
$50,000 - $74,999	6	17.1	8	23.5	14	20.3	
$75,000 - $99,999	4	11.4	9	26.5	13	18.8	
$100,000 - $200,000	-	-	5	14.7	5	7.2	
More than $200,000	-	-	3	8.8	3	4.3	
Age							.036
18 – 24	11	31.4	2	5.9	13	18.8	
25 – 34	13	37.1	15	44.1	28	40.6	
35 – 44	7	20	6	17.6	13	18.8	
45 – 54	3	8.6	8	23.5	11	15.9	
55 – 64	1	2.9	1	2.9	2	2.9	
65 – 74	-	-	2	5.9	2	2.9	

Demographics and Pre-Screen measures largely remained unaffected with two exceptions. In the supplementary analysis, the denial subscale scores were not significantly different between the two groups. Additionally, there were significant differences detected in age between the CPTSD group and the Control group, with the CPTSD group seemingly consisting of a younger age bracket than the Control group.

Appendix E: Supplementary Analysis with Positive Denial Subscale Score Removal

This supplementary analysis duplicates the results from the main analysis with the caveat that participants who reported a positive denial subscale score (score > 0) were removed from the analysis (n = 11). Similar to the previous supplementary studies, trends observed in the main analyses were present within this supplementary analysis as well, suggesting robustness. There were additional significant interactions observed within this supplementary analysis, notably a Group x Phase x Agent interaction for prediction accuracy. Data analysis used in the supplementary analysis were identical to that of the main analysis.

Table 1

ANOVA Results for Impression

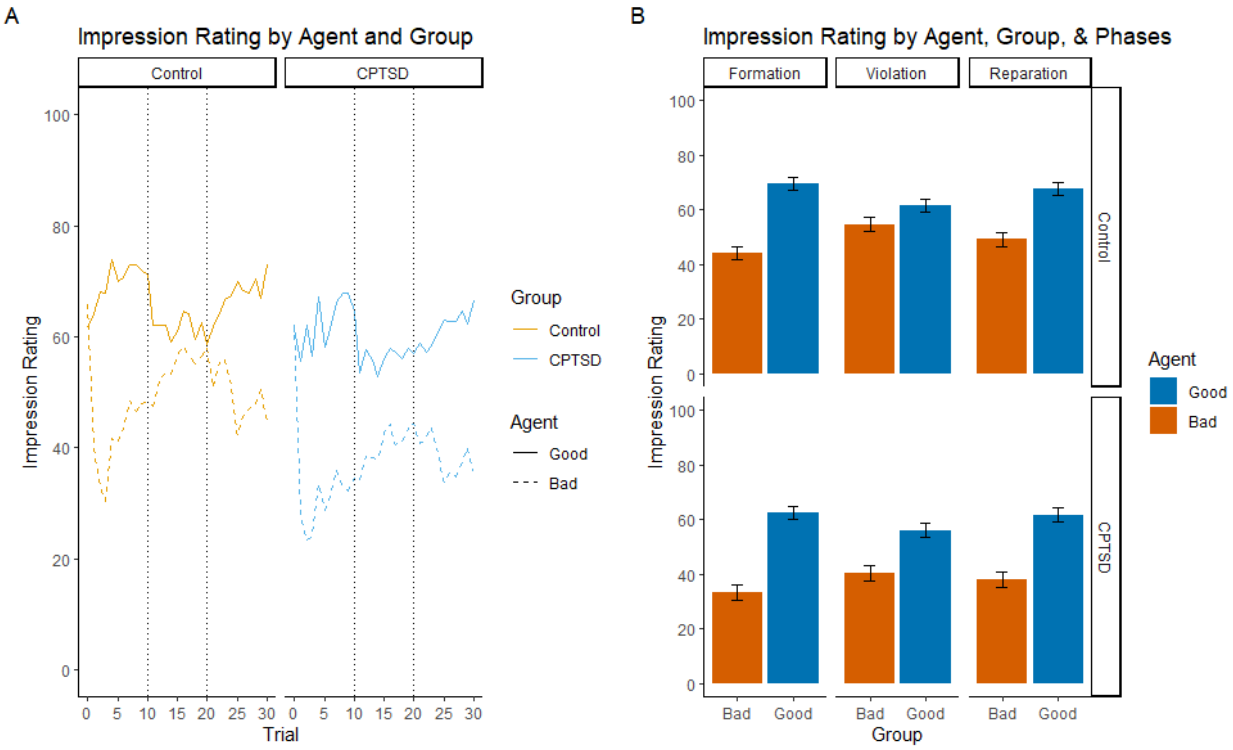
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 59	3.57	.057	.064	.
Phase	1.62, 95.6	1.62	.027	.207	
Group : Phase	1.62, 95.6	0.268	.005	.718	
Agent	1, 59	76.3	.564	<.001	***
Group : Agent	1, 59	1.72	.028	.195	
Phase : Agent	1.96, 115.4	30.6	.341	<.001	***
Group : Phase : Agent	1.96, 115.4	0.73	.012	.481	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes ($.01 < \eta_p^2 < .06$) for Group, Phase, Group x Agent, and Group x Agent x Phase. Large effects ($\eta_p^2 > .14$) were present for Phase and Agent x Phase.

ANOVA results are consistent with main analyses, with the exception that there is no longer a marginally significant effect observed for Phase in the supplementary analysis.

Figure 1

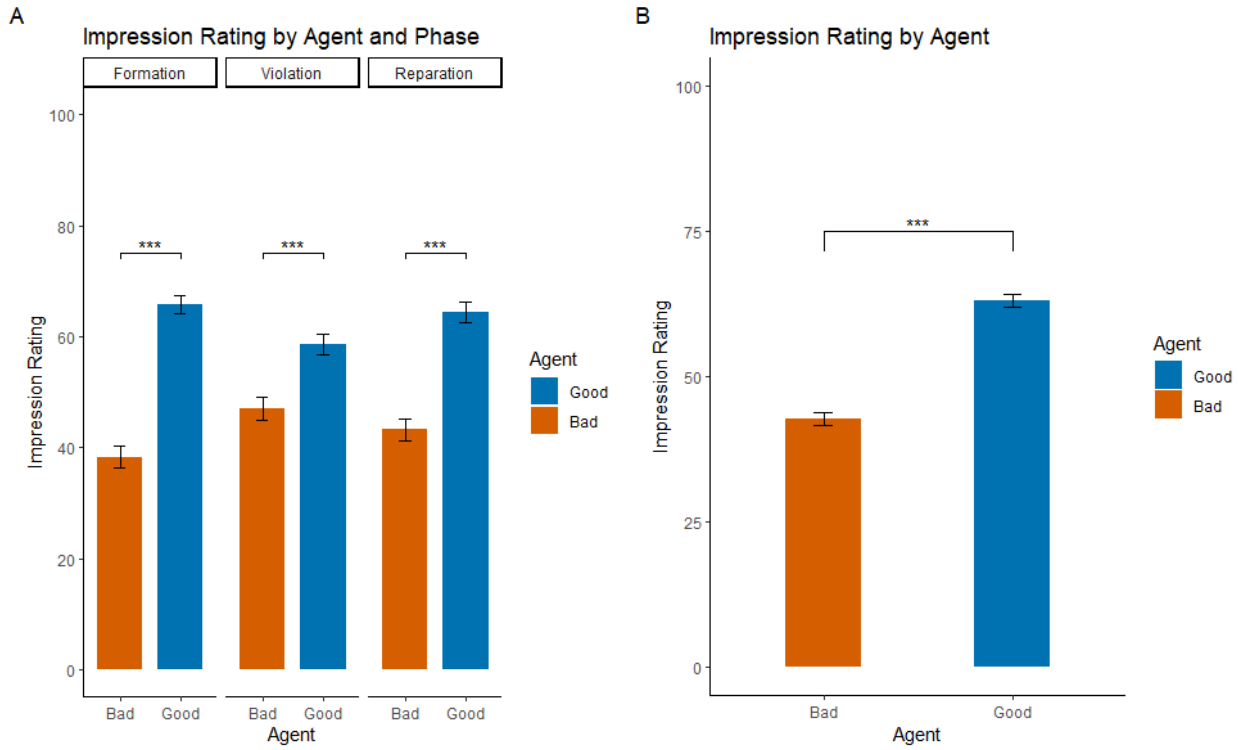
Impression Ratings by Agent and Group throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Figure 2

Significant Main Effects and Interactions for Impression Rating



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 2

ART ANOVA Results for Certainty

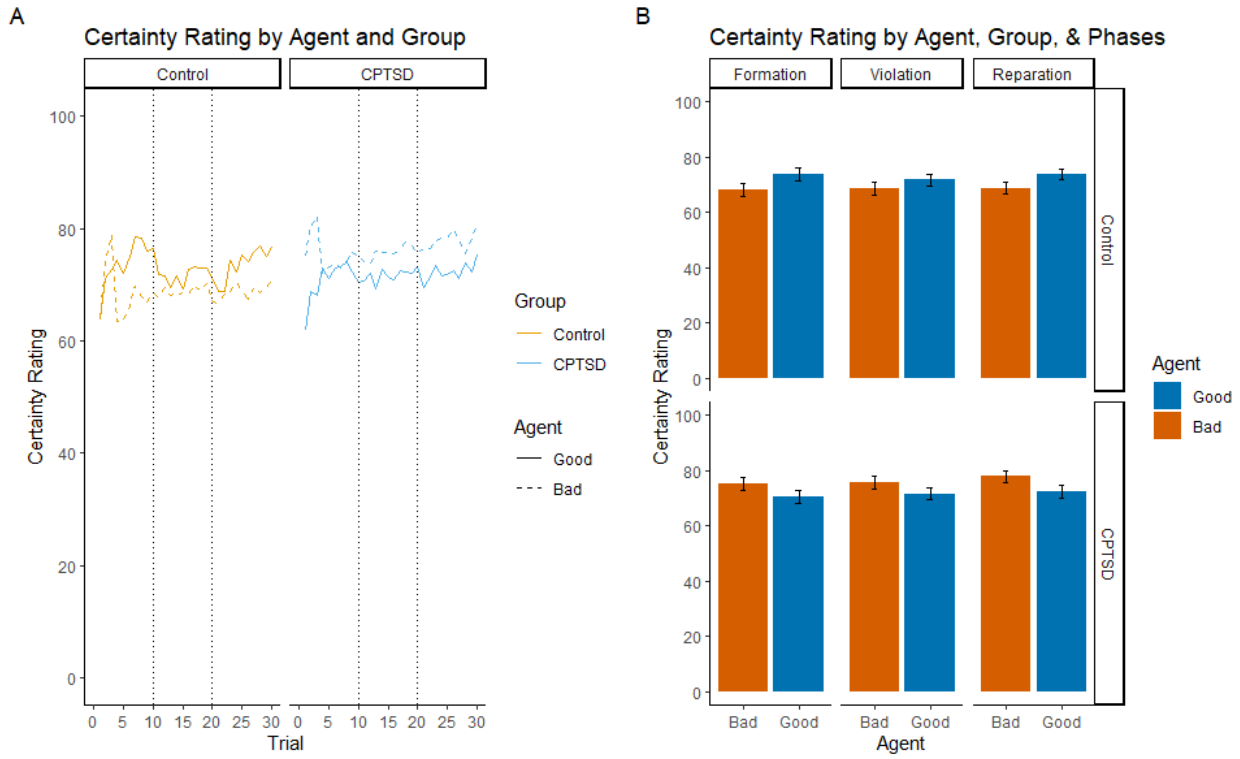
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 59	0.576	.01	.451	
Phase	2, 295	0.610	.004	.544	
Group : Phase	2, 295	0.118	.001	.889	
Agent	1, 295	0.671	.002	.414	
Group : Agent	1, 295	20.7	.065	<.001	***
Phase : Agent	2, 295	0.063	<.001	.94	
Group : Phase : Agent	2, 295	0.450	.003	.638	

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) indicated a small effect size ($.01 < \eta_p^2 < .06$) for Group x Agent.

Trends observed in the main analysis were replicated for certainty ratings in the supplementary analysis, with no notable differences.

Figure 3

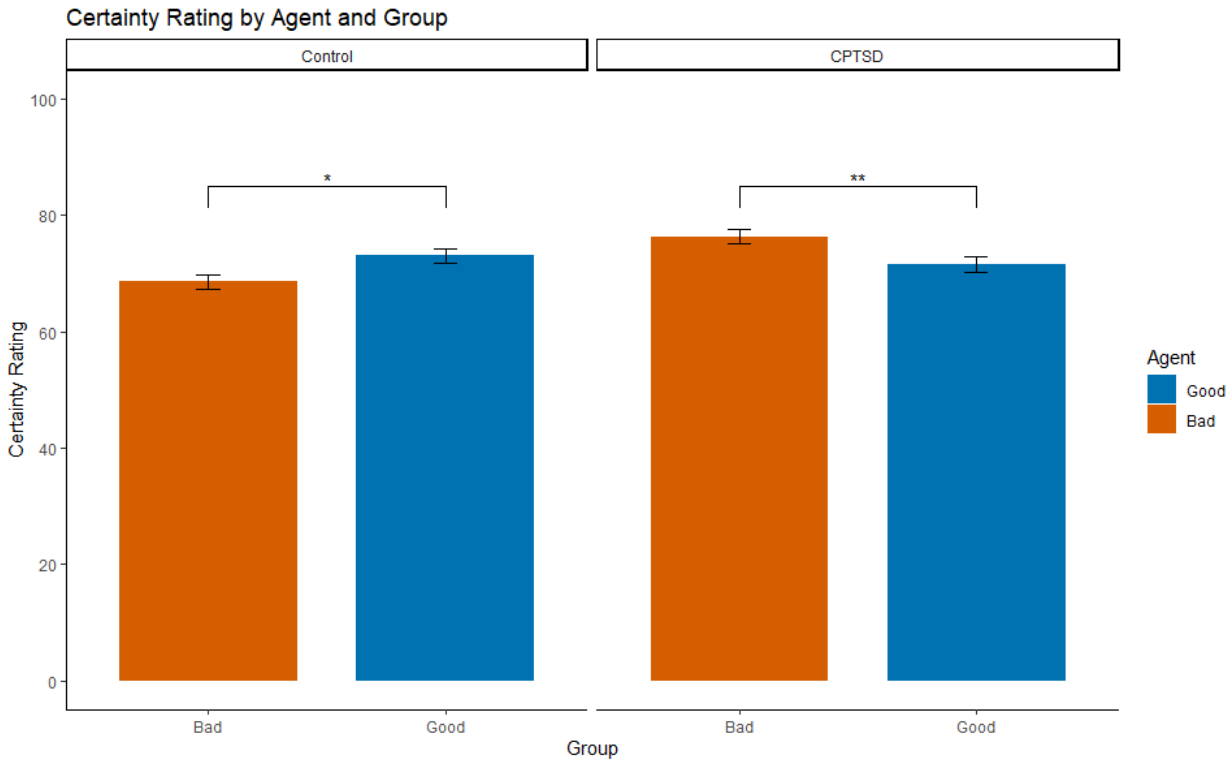
Certainty Ratings by Group and Agent throughout Task



Note. (A) Dotted lines indicate the beginning and endings of various task phases. Trials reflect the 30 instances in which participants were asked to rate the agents' moral character. (B) Error bars represent 95% CI.

Figure 4

Significant Main Interaction for Certainty Ratings



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 3

ART ANOVA Results for Prediction Accuracy

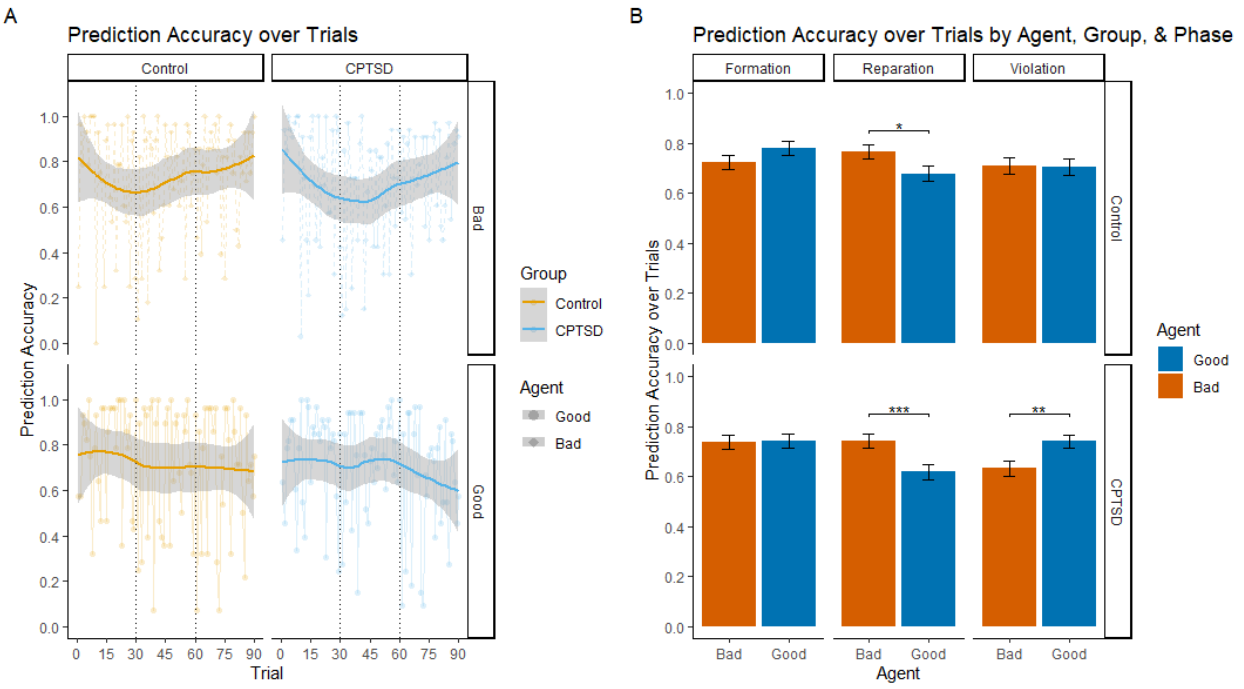
Measure	*df*	*F*	η_p^2	*p* value	
Group	1, 59	1.95	.032	.168	
Phase	2, 295	10.6	.067	<.001	***
Group : Phase	2, 295	0.299	.002	.742	
Agent	1, 295	0.611	.002	.435	
Group : Agent	1, 295	0.058	<.001	.81	
Phase : Agent	2, 295	26.9	.154	<.001	***
Group : Phase : Agent	2, 295	5.95	.039	.003	**

Note. · p<0.1, *p<.05, **p<.01, ***p<.001. Partial Eta Squared (η_p^2) generally indicated small effect sizes (.01 < η_p^2 < .06) for Group, Phase, and Group x Phase x Agent. Large effects (η_p^2 <.14) were present for Phase x Agent.

Results in the supplementary analysis duplicated that of the main analysis with the exception of the fact that a significant interaction for Group x Phase x Agent was observed in the supplementary analysis for prediction accuracy, $F(2, 295) = 5.95$, $p = .003$, $\eta_p^2 = .039$.

Figure 5

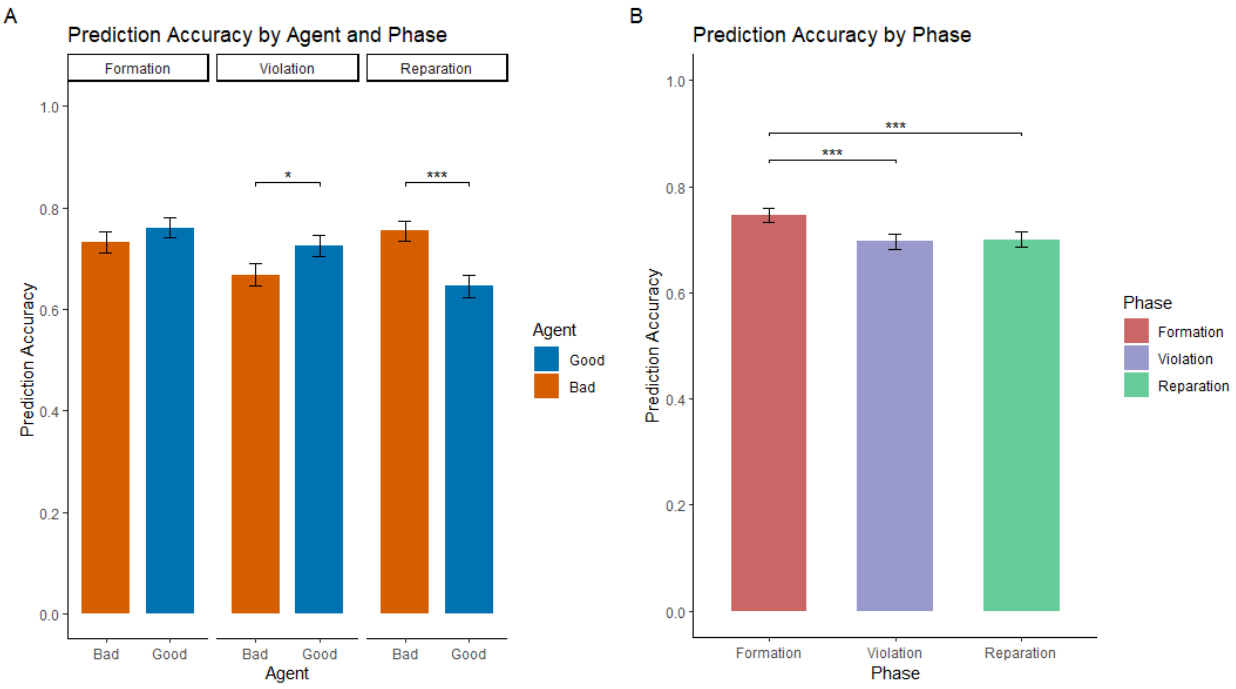
Prediction Accuracy by Group and Agent throughout Task, including Significance



Note. *p<.05, **p<.01, ***p<.001. (A) Dotted lines indicate the beginning and endings of various task phases. Average Prediction Accuracy per Trial was mapped in the background and a local polynomial regression (solid line) was fitted over it. Error bands represent 95% CI. (B) Error bars represent 95% CI.

Figure 6

Significant Main Effects and Interactions for Prediction Accuracy



Note. *p<.05, **p<.01, ***p<.001. Error bars represent 95% CI.

Table 4

Pre-Screen Trauma Surveys

	CPTSD (n = 33, 54.1%)	Control (n = 28, 45.9%)	Overall (n = 61)	*U*	*p* value	r
ACEs	5	0	3	-6.14	<.001	.786
CTQ	70	40.5	48	-6.41	<.001	.821
Emotional Abuse	18	6	10.5	-6.20	<.001	.794
Emotional Neglect	19	8	13.5	-6.23	<.001	.797
Physical Abuse	9	5	6	-5.01	<.001	.641
Physical Neglect	11	5	7.5	-6.02	<.001	.770
Sexual Abuse	11	5	5	-4.86	<.001	.622

Note. Values reflect medians due to non-parametric distribution. *p* values were derived from Mann Whitney U Tests that were conducted to compare CPTSD and Control groups. Wilcoxon effect sizes (r) in these cases exhibit large effects ($r \geq .5$).

Table 5

Participant Demographics

	CPTSD (n = 33, 54.1%)		Control (n = 28, 45.9%)		Overall (n = 61)		*p* value
	n	%	n	%	n	%	
Race							.071
White	24	72.7	22	78.6	46	75.4	
Asian	1	3	5	17.9	6	9.8	
American Indian or Alaskan Native	2	6.1	-	-	2	3.3	
Multiple Selected	5	15.2	1	3.6	6	9.8	
Other	1	3	-	-	1	1.6	
Education							.676
High school graduate or equivalent	4	12.1	2	7.1	6	9.8	
Some college but not degree	11	33.3	7	25	18	29.5	
Associate degree	3	9.1	1	3.6	4	6.6	
Bachelor's degree	11	33.3	11	39.3	22	36.1	
Master's degree	4	12.1	6	21.4	10	16.4	
Doctoral degree	-	-	1	3.6	1	1.6	
Professional degree	-	-	-	-	-	-	
Gender							.378
Female	22	66.7	17	60.7	39	63.9	
Male	8	24.2	11	39.3	18	31.1	
Gender Nonbinary	2	6.1	-	-	2	3.3	
Other	1	3	-	-	1	1.6	
Relationship Status							.551

Never Married	25	75.8	17	60.7	42	68.9	
Married	7	21.2	9	32.1	16	26.2	
Divorced	1	3	1	3.6	2	3.3	
Widowed	-	-	1	3.6	1	1.6	
Financial Status							<.001
Less than $25,000	15	45.5	1	3.6	16	26.2	
$25,000 - $49,999	9	27.3	7	25	16	26.2	
$50,000 - $74,999	6	18.2	8	28.6	14	23	
$75,000 - $99,999	3	9.1	5	17.9	8	13.1	
$100,000 - $200,000	-	-	4	14.3	4	6.6	
More than $200,000	-	-	3	10.7	3	4.9	
Age							.154
18 – 24	10	30.3	2	7.1	12	19.7	
25 – 34	12	36.4	12	42.9	24	39.3	
35 – 44	7	12.1	6	21.4	13	21.3	
45 – 54	3	9.1	6	21.4	9	14.8	
55 – 64	1	3	1	3.6	2	3.3	
65 – 74	-	-	1	3.6	1	1.6	

In this supplementary analysis, no major differences in demographics and prescreen measures were observed. Results were virtually identical to that of the main analysis. Notably, in this analysis, as only participants with a denial score of 0 were included, statistical testing was not conducted for that subscale.

Appendix F: Forms

Form for Collaboration in Senior Thesis Work

Please use this form to indicate the relationship between previous work and your senior thesis and to indicate whether your thesis involved collaboration with others.

Indicate below whether there is any overlap between your senior thesis and earlier work that you did for junior reports, junior papers, or papers for various courses.

Overlap x

No Overlap _____

If you checked the box indicating that there is overlap between your senior thesis and previous work, please describe the overlap on a **separate page**, and include it within the thesis after this form.

Readers of your thesis may, if they choose, ask to see earlier papers that you indicate have some overlap with your senior thesis.

Indicate below whether all or part of your thesis resulted from work done collaboratively with one or more other people.

Collaboration _____

No Collaboration x

If you checked the box indicating that your thesis work was done entirely, or in part, in collaboration with other people, describe the nature of the collaboration and what resulted from it on a separate page, and include it within the thesis after this form.

Statement of Overlap

This paper draws heavily on my spring junior paper, which proposed a research project that was then conducted within this senior thesis. As such, this senior thesis includes a lot of the same material that was utilized within the literature review, the objectives and hypotheses, as well as the methodology section of the original junior paper. All sections were reworked to include new information and additional modifications but much of the original language used was retained. The data-analysis, results, and discussion section all represent new work that was conducted this year.

Approval Form for Undergraduate Research Involving Experimental Animals

All research involving experimental animals at Princeton University must receive the prior approval from the Institutional Animal Care and Use Committee (IACUC). The IACUC bases decision about approval on the NRC Guide for the Care and Use of Laboratory Animals. All students conducting research involving animals as part of their junior independent work or senior thesis must receive approval from the IACUC prior to beginning their research. Students should consult first with their advisers about whether the procedures they intend to use are already covered by previously approved submission to the IACUC. The IACUC meets only once a month and it is common for new submissions to require revision before receiving approval so students are strongly encouraged to attend to IACUC issues early in their planning.

Did your Senior Thesis research involve the use of experimental animals?

Yes _____ No x

If you answered "Yes" to the above, you *must* also include a statement at the beginning of your methods section that verifies the work you have done with animals was approved by the Princeton University IACUC.

Lastly, please include this form at the back of your thesis (even if you answered "No"). **If you answered "Yes," please record the IACUC protocol number and date, below.**

IACUC # _____ Approval Date _____

Approval Form for Undergraduate Research Involving Human Subjects

The Institutional Review Board for Human Subjects (IRB) is charged by the University Research Board with the task of protecting the interests and rights of human subjects involved in Princeton research. The IRB's responsibility includes the oversight of research conducted by undergraduate as part of their junior independent work and senior thesis work as well as that conducted in fulfillment of course requirements. All students conducting research involving human subjects as part of their junior independent work or senior thesis must receive approval from the IRB prior to beginning their research. Obviously, the sooner students submit their requests to IRB the sooner they will receive this approval. Students should be encouraged to submit their materials to the IRB as soon as possible in the semester. The IRB meets only once a month and it is common for student submissions to require revisions, primarily because of the incompleteness of the original submission, before receiving approval.

Did your Senior Thesis involve research with human subjects?

Yes ___x__ No _____

If your Senior Thesis DID involve research with human subjects, please indicate your IRB Case Number below.

Case # 11968 Approval Date 12 /16/2019